```
                  GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

M nucleic - nucleic search, using sw model

un on:          October 22, 2003, 07:27:11 ; Search time 5513 Seconds
                                             (without alignments)
                                             11197.647 Million cell updates/sec

ıtle:           US-10-006-852-1
erfect score:   1509
equence:        1 atggtgctctcccacgccgt..........agacgagtggtatctgctaa 1509

coring table:   IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

earched:        2888711 seqs, 20454813386 residues

otal number of hits satisfying chosen parameters:      5777422

ınimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

atabase :        GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*

     Pred. No. is the number of results predicted by chance to have a
```

```
          score greater than or equal to the score of the result being printed,
          and is derived by analysis of the total score distribution.

                                        SUMMARIES
                         %
 Result                Query
   No.     Score       Match Length DB  ID                         Description
 ------------------------------------------------------------------------------
       1    1509   100.0    1509  8  ATU10034                     U10034 Arabidopsis
       2  1507.4    99.9    1509  6  AX507115                     AX507115 Sequence
       3  1507.4    99.9    1509  6  AX651552                     AX651552 Sequence
       4  1507.4    99.9    1509  8  BT001047                     BT001047 Arabidops
       5  1507.4    99.9    1813  8  AY094464                     AY094464 Arabidops
       6   907.4    60.1    1776  8  AF353615                     AF353615 Nicotiana
       7   883.4    58.5    1705  8  AF020425                     AF020425 Nicotiana
       8   881.8    58.4    1785  8  PETGADX                      L16797 Petunia hyb
       9   881.8    58.4    1785  8  PETGLUDECA                   L16977 Petunia hyb
      10   881.8    58.4    1929  8  NTU54774                     U54774 Nicotiana t
      11     877    58.1    1672  8  AF352732                     AF352732 Nicotiana
      12   870.2    57.7    1485  8  AY124873                     AY124873 Arabidops
      13   870.2    57.7    1676  8  ATU46665                     U46665 Arabidopsis
      14   870.2    57.7    1681  8  ATU49937                     U49937 Arabidopsis
      15   870.2    57.7    1695  8  AF428372                     AF428372 Arabidops
      16   868.6    57.6    1685  8  AF428294                     AF428294 Arabidops
      17   864.2    57.3    1745  8  AF506366                     AF506366 Nicotiana
      18   861.4    57.1    1650  8  AY081259                     AY081259 Arabidops
      19   857.8    56.8    1771  8  AF020424                     AF020424 Nicotiana
      20   854.2    56.6    1482  6  AX412801                     AX412801 Sequence
      21   854.2    56.6    1482  6  AX506242                     AX506242 Sequence
      22   854.2    56.6    1482  6  AX651853                     AX651853 Sequence
      23   854.2    56.6    1482  8  AY124860                     AY124860 Arabidops
      24   854.2    56.6    1792  8  AF361836                     AF361836 Arabidops
      25   838.2    55.5    1712  8  AK118125                     AK118125 Arabidops
      26   834.2    55.3    1479  6  AX654603                     AX654603 Sequence
      27   834.2    55.3    1502  8  AY187941                     AY187941 Oryza sat
      28     770    51.0    1605  8  AY084689                     AY084689 Arabidops
      29     769    51.0    1779  8  AB056060                     AB056060 Oryza sat
      30   736.6    48.8    1783  6  A39964                       A39964 Sequence 8
      31   736.6    48.8    1783  6  AR070945                     AR070945 Sequence
      32   736.6    48.8    1783  8  LEGDL                        X80840 L.esculentu
      33   711.4    47.1    1180  6  BD017613                     BD017613 Method fo
      34   709.6    47.0    1455  6  AX653206                     AX653206 Sequence
      35   670.6    44.4    1781  8  AB056061                     AB056061 Oryza sat
      36   560.2    37.1    1467  6  AX653024                     AX653024 Sequence
      37   487.6    32.3   75188  8  AB005238                     AB005238 Arabidops
 c    38   432.4    28.7  136098  8  AC137696                     AC137696 Genomic s
      39   415.2    27.5    3983  8  AB056063                     AB056063 Oryza sat
 c    40   415.2    27.5  147227  8  OSJN00013                    AL606448 Oryza sat
      41     393    26.0    5160  6  AX461514                     AX461514 Sequence
      42     393    26.0   81490  8  F12P19                       AC009513 Sequence
      43   386.2    25.6    1545  6  BD008061                     BD008061 Glutamate
 c    44   385.8    25.6   76976  8  AC006532                     AC006532 Arabidops
      45   380.4    25.2     906  8  AF358767                     AF358767 Oryza sat


                                        ALIGNMENTS


RESULT 1
ATU10034
LOCUS       ATU10034                1509 bp    mRNA    linear   PLN 30-MAR-1996
DEFINITION  Arabidopsis thaliana Columbia glutamate decarboxylase (GAD) mRNA,
            complete cds.
ACCESSION   U10034
VERSION     U10034.1  GI:497978
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1509)
  AUTHORS   Arazi,T., Baum,G., Snedden,W.A., Shelp,B.J. and Fromm,H.
```

```
 TITLE      Molecular and biochemical analysis of calmodulin interactions with
            the calmodulin-binding domain of plant glutamate decarboxylase
 JOURNAL    Plant Physiol. 108 (2), 551-561 (1995)
 MEDLINE    95334488
  PUBMED    7610159
FERENCE     2  (bases 1 to 1509)
 AUTHORS    Fromm,H.
 TITLE      Direct Submission
 JOURNAL    Submitted (25-MAY-1994) Hillel Fromm, Department of Plant Genetics,
            Weizmann Institute of Science, Rehovot 76100, Israel
ATURES               Location/Qualifiers
    source           1. .1509
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="35B3T7"
    gene             1. .1509
                     /gene="GAD"
    CDS              1. .1509
                     /gene="GAD"
                     /EC_number="4.1.1.15"
                     /note="GAD1 isoform; similar to the GAD2 isoform of
                     Arabidopsis thaliana, GenBank Accession Number U49937"
                     /codon_start=1
                     /product="glutamate decarboxylase"
                     /protein_id="AAA93132.1"
                     /db_xref="GI:497979"
                     /translation="MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEF
                     EDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEIS
                     DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSR
                     VIHKISLGQEKSESNSDNLMVTVKKSDIDKQRDIITGWKKFVADRKKTSGIC"
    misc_feature     1417. .1506
                     /gene="GAD"
                     /note="calmodulin-binding domain"
ASE COUNT      433 a    321 c    385 g    370 t
IGIN

 Query Match             100.0%;  Score 1509;  DB 8;  Length 1509;
 Best Local Similarity   100.0%;  Pred. No. 0;
 Matches 1509;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

y         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

y        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

y        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

y        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

y        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

y        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Qy       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
```

```
       1501 ATCTGCTAA 1509
            |||||||||
       1501 ATCTGCTAA 1509


ESULT 2
X507115
OCUS       AX507115                  1509 bp    DNA     linear   PAT 27-SEP-2002
EFINITION  Sequence 1810 from Patent WO0216655.
CCESSION   AX507115
ERSION     AX507115.1  GI:23388352
EYWORDS    .
OURCE      Arabidopsis thaliana (thale cress)
 ORGANISM  Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
EFERENCE   1
 AUTHORS   Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
 TITLE     Stress-regulated genes of plants, transgenic plants containing
           same, and methods of use
 JOURNAL   Patent: WO 0216655-A 1810 28-FEB-2002;
           The Scripps Research Institute (US) ; Syngenta Participations AG
           (CH)
EATURES             Location/Qualifiers
    source          1. .1509
                    /organism="Arabidopsis thaliana"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:3702"
ASE COUNT      432 a    322 c    385 g    370 t
RIGIN

 Query Match              99.9%;  Score 1507.4;  DB 6;  Length 1509;
 Best Local Similarity    99.9%;  Pred. No. 0;
 Matches 1508;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy          541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy          601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
                |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db          601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy          661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy          721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy          781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy          841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy          901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy          961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy         1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy         1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy         1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy         1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy         1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy         1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Qy         1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy         1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

Qy         1501 ATCTGCTAA 1509
                |||||||||
Db         1501 ATCTGCTAA 1509


RESULT 3
AX651552
LOCUS       AX651552                1509 bp    DNA     linear   PAT 22-MAR-2003
DEFINITION  Sequence 360 from Patent WO03000898.
```

```
CESSION   AX651552
RSION     AX651552.1  GI:29154370
YWORDS    .
URCE      Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
FERENCE   1
AUTHORS   Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
          Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
TITLE     Plant genes involved in defense against pathogens
JOURNAL   Patent: WO 03000898-A 360 03-JAN-2003;
          Syngenta Participations AG (CH)
ATURES             Location/Qualifiers
   source          1. .1509
                   /organism="Arabidopsis thaliana"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:3702"
SE COUNT       432 a    322 c    385 g    370 t
IGIN

Query Match             99.9%;  Score 1507.4;  DB 6;  Length 1509;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 1508;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
            |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
        601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
```

```
Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Qy       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

Qy       1501 ATCTGCTAA 1509
              |||||||||
Db       1501 ATCTGCTAA 1509


RESULT 4
BT001047
LOCUS       BT001047                1509 bp    mRNA    linear   PLN 18-OCT-2002
DEFINITION  Arabidopsis thaliana At5g17330/MKP11_18 mRNA, complete cds.
ACCESSION   BT001047
VERSION     BT001047.1  GI:24111354
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1509)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
```

```
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
            Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
TITLE       Arabidopsis ORF clones
JOURNAL     Unpublished
FERENCE     2  (bases 1 to 1509)
AUTHORS     Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
            Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yamada,K.,
            Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
            Ecker,J.R.
TITLE       Direct Submission
JOURNAL     Submitted (18-OCT-2002) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
MMENT       RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
            Chen,H., Kim,C.J., Shinn,P., Banh,J. Bowser,L., Chan,M.M.,
            Chang,E., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S.,
            Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M.,
            Wu,H.C., Yamada,K., Yamamura, Y., Yu,G., Davis,R.W.,  Theologis,A.,
            and Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
ATURES               Location/Qualifiers
  source             1. .1509
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="U21967"
                     /note="This clone is in pUNI 51
                     ecotype: Columbia"
   CDS               1. .1509
                     /note="glutamate decarboxylase 1 (GAD 1) (sp|Q42521)"
                     /codon_start=1
                     /product="At5g17330/MKP11_18"
                     /protein_id="AAN46801.1"
                     /db_xref="GI:24111355"
                     /translation="MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVAAILGSTLNGEF
                     EDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEIS
                     DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSR
                     VIHKISLGQEKSESNSDNLMVTVKKSDIDKQRDIITGWKKFVADRKKTSGIC"
ASE COUNT      432 a    322 c    385 g    370 t
RIGIN

 Query Match             99.9%;  Score 1507.4;  DB 8;  Length 1509;
 Best Local Similarity   99.9%;  Pred. No. 0;
 Matches 1508;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

y           1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

Qy        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
```

```
      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

      1501 ATCTGCTAA 1509
           |||||||||
      1501 ATCTGCTAA 1509


SULT 5
094464
CUS         AY094464                1813 bp    mRNA    linear   PLN 05-MAY-2002
FINITION    Arabidopsis thaliana AT5g17330/MKP11_18 mRNA, complete cds.
CESSION     AY094464
RSION       AY094464.1  GI:20453186
YWORDS      FLI_CDNA.
URCE        Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
FERENCE     1  (bases 1 to 1813)
AUTHORS     Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
            Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
            Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
            Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
TITLE       Arabidopsis cDNA clones
JOURNAL     Unpublished
FERENCE     2  (bases 1 to 1813)
AUTHORS     Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
            Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
            Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
            Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
TITLE       Direct Submission
JOURNAL     Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
MMENT       RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.
```

```
            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H.,
            Cheuk,R., Kim,C.J.,  Meyers,M.C., Banh,J. Bowser,L., Chang,E.,
            Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B.,
            Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
            Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W.,  Theologis,A.,
            and Ecker,J.R.

            Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1813
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="RAFL09-48-C10(R21967)"
                     /note="ecotype: Columbia"
     CDS             74. .1582
                     /note="glutamate decarboxylase 1 (GAD 1) (sp|Q42521)"
                     /codon_start=1
                     /product="AT5g17330/MKP11_18"
                     /protein_id="AAM19834.1"
                     /db_xref="GI:20453187"
                     /translation="MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVAAILGSTLNGEF
                     EDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEIS
                     DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSR
                     VIHKISLGQEKSESNSDNLMVTVKKSDIDKQRDIITGWKKFVADRKKTSGIC"
BASE COUNT      504 a    384 c    417 g    508 t
ORIGIN

  Query Match             99.9%;  Score 1507.4;  DB 8;  Length 1813;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1508;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 133

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 193

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 253

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 313

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 373

Qy        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 433

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 493

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 553
```

```
)y      481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       554 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 613

y       541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       614 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 673

y       601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
            |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
b       674 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 733

y       661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       734 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 793

y       721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       794 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 853

y       781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       854 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 913

y       841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       914 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 973

y       901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       974 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 1033

y       961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      1034 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1093

y      1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)      1094 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1153

y      1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)      1154 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1213

'      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)      1214 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1273

'      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)      1274 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1333

'      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)      1334 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1393

'      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)      1394 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1453

'      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)      1454 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1513

'      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)      1514 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1573

'      1501 ATCTGCTAA 1509
            |||||||||
       1574 ATCTGCTAA 1582
```

```
RESULT 6
AF353615
LOCUS       AF353615                1776 bp    mRNA    linear   PLN 02-MAY-2003
DEFINITION  Nicotiana tabacum glutamate decarboxylase isozyme 3 mRNA, complete
            cds.
ACCESSION   AF353615
VERSION     AF353615.1  GI:13752461
KEYWORDS    .
SOURCE      Nicotiana tabacum (common tobacco)
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1776)
  AUTHORS   Yevtushenko,D., McLean,M.D., Peiris,S.E., Van Cauwenberghe,O.R. and
            Shelp,B.J.
  TITLE     Two isoforms of tobacco glutamate decarboxylase are regulated by
            calcium/calmodulin and differ in organ distribution
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1776)
  AUTHORS   McLean,M.D., Yevtushenko,D. and Shelp,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-2001) Department of Plant Agriculture, University
            of Guelph, Division of Biotechnology, Bovey Building, Guelph,
            Ontario N1G 2W1, Canada
FEATURES             Location/Qualifiers
     source          1. .1776
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4097"
     CDS             89. .1564
                     /EC_number="4.1.1.15"
                     /function="converts glutamate to gamma-aminobutyric acid"
                     /note="GAD; GAD4; NtGAD4"
                     /codon_start=1
                     /product="glutamate decarboxylase isozyme 3"
                     /protein_id="AAK38667.1"
                     /db_xref="GI:13752462"
                     /translation="MVLSKTSSESDVSVHSTFASRYVRTSLPRFEMAENSIPKEAAFQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIARLFNAPLEEKETAVGVGTVGSSEAIMLAGLAFKRNWQNKRKAEGKPYNKPNIVTG
                     ANVQVCWEKFANYFEVELKEVKLREGYYVMDPVQAVEMVDENTICVAAILGSTLNGEF
                     EDVKLLNDLLIEKNKQTGWNTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWVIWRTKQDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGYEGYRNVMENCRENAIVLREGLEKTGRFNIVSKDEGVPLVAFSLKDNSRHNEFEVS
                     ETLRRFGWIVPAYTMPADAQHVTVLRVVIREDFSRTLAERLVLDIVKVLHELDTLPAR
                     LSAKLEEVKLVKNGKKFELEVQREVTNYWKKFVLARKAPVC"
BASE COUNT      520 a    343 c    418 g    491 t      4 others
ORIGIN

  Query Match             60.1%;  Score 907.4;  DB 8;  Length 1776;
  Best Local Similarity   79.0%;  Pred. No. 7.5e-254;
  Matches 1079;  Conservative    0;  Mismatches  286;  Indels    0;  Gaps    0;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||| |||||  |  ||   || ||||||||||| || || |||||||| || || ||
Db         89 ATGGTTCTCTCGAAAACCTCCTCTGAGTCGGACGTTTCGGTACACTCCACTTTTGCCTCT 148

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || || || || ||||| ||||| |||||  ||||| |||| || ||||| || || ||
Db        149 CGCTATGTTCGAACTTCCCTTCCAAGGTTTGAGATGGCGGAGAATTCGATACCAAAAGAG 208

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              ||||| | ||| || || ||||||||| |||||||||||||||||||| || |||||||
Db        209 GCGGCATTTCAAATAATTAACGACGAGTTGATGCTTGACGGGAATCCAAGGCTGAACTTG 268

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              || || |||||||| |||||||||||||| ||||||||||| || || |||  ||||||
Db        269 GCTTCATTTGTGACAACATGGATGGAGCCAGAGTGTGATAAGCTTATGATGGACTCCATT 328

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||||||||||||||| || ||||| || ||||| || |||||||| || ||
```

```
 329 AACAAGAACTATGTTGACATGGATGAATACCCTGTTACCACTGAGCTTCAGAATCGCTGC 388

 301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
     ||||||||||| |||| | ||||||| || ||  ||||||||  |||||| || || |||
 389 GTGAACATGATAGCACGTTTATTCAACGCGCCACTAGAAGAGAAGGAGACAGCAGTTGGA 448

 361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
     || || || ||||| |||||||||||||||||| | || ||  | || |||||| | ||
 449 GTGGGTACAGTTGGTTCATCGGAGGCCATAATGCTAGCGGGGCTTGCATTCAAGAGAAAT 508

 421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
     ||||| ||||| ||||||||||| ||||||||||   | |||| ||||||||||||||| ||
 509 TGGCAAAACAAACGCAAAGCTGAGGGCAAACCTTACAATAAGCCCAACATTGTCACTGGC 568

 481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
     ||||||||||| ||||| ||||||||||||| |  || ||||| || ||| | || |||
 569 GCCAATGTTCAGGTGTGCTGGGAGAAATTTGCCAACTATTTTGAAGTGGAATTGAAAGAA 628

 541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
     || ||  | || ||||| ||||||||||||||||||    || ||||| || |||||||||
 629 GTAAAGCTAAGGGAAGGGTACTATGTGATGGACCCAGTCCAGGCTGTGGAGATGGTTGAT 688

 601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
     |||||||||||||||||||| |  ||  | ||||| || ||||||||||||| ||||||
 689 GAGAACACCATTTGTGTTGCTGCAATCTTGGGTTCAACCCTTAATGGAGAATTTGAAGAT 748

 661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
     || || |||||||| ||||| ||| | ||||||||||||  |||| |||||| | ||||||
 749 GTCAAGCTCTTGAATGATCTTTTGATTGAAAAGAACAAGCAAACTGGATGGAACACACCA 808

 721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
     || || |||||||| ||||||||| |||||||||||||||| ||  |||| || ||  ||||
 809 ATTCATGTGGATGCAGCAAGTGGTGGATTCATTGCACCATTCCTGTACCCAGAGCTGGAG 868

 781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
     ||||||||||| |||||||| |||||||| || |||||||||||| ||||| ||||| |||
 869 TGGGACTTTAGGCTTCCCTTAGTGAAGAGCATAAATGTGAGTGGGCACAAATATGGGCTT 928

 841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
     || || || || ||||||||||| || ||||| | |||| | |||||||||||| ||||||
 929 GTCTATGCTGGTATTGGTTGGGTTATTTGGAGGACCAAACAAGATTTGCCTGAAGAACTC 988

 901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
     || ||||| ||||| ||||||||| ||||| || ||||||||||||||||||||||||||
 989 ATTTTCCACATCAACTATCTTGGAGCTGATCAGCCCACCTTTACTCTCAATTTCTCCAAA 1048

 961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
     |||||||||||||||||||||||||| || || || ||||| |||||| | ||||| |||
1049 GGTTCAAGTCAAGTCATTGCTCAATATTATCAGCTAATCCGTTTGGGCTATGAGGGGTAC 1108

1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
      ||||||| |||||||| ||  | || |||   || || ||||| |||||||| || ||
1109 CGAAATGTAATGGAGAACTGTCGCGAAAATGCCATTGTGCTAAGAGAAGGACTCGAAAAA 1168

1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
     |||| | | || ||||| |||| || || || || || ||  | || ||| || ||| |
1169 ACAGGACGTTTTAACATAGTCTCCAAAGATGAAGGTGTCCCTTTGGTGGCCTTTTCCCTC 1228

1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
     || || |  ||| |||||| |||||||||  | ||||| | ||| || ||||  || |||
1229 AAGGACAATAGCCGTCACAATGAGTTCGAGGTGTCCGAGACGCTCCGTAGGTTCGGGTGG 1288

1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
     || || ||||||||||| |||||  |  | || |||||| |||| || ||||||||||
1289 ATCGTCCCGGCCTACACGATGCCCGCTGACGCCCAACACGTCACGGTGCTTCGTGTGGTG 1348

1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
     ||| | || || |||||| |||| || || ||| | || ||  |||| || |  || ||
1349 ATCCGGGAGGACTTCTCGCGAACCCTAGCAGAGCGTCTCGTCCTCGACATTGTCAAGGTC 1408

1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATA 1365
      | |  ||||| ||    |||||  | ||  |||    |||| ||
1409 CTCCACGAGCTGGACACACTTCCAGCTAGGCTGAGCGCCAAATTA 1453
```

```
RESULT 7
AF020425
LOCUS       AF020425                1705 bp    mRNA    linear   PLN 24-JUN-1998
DEFINITION  Nicotiana tabacum glutamate decarboxylase isozyme 1 (NtGAD1) mRNA,
            complete cds.
ACCESSION   AF020425
VERSION     AF020425.1  GI:3252855
KEYWORDS    .
SOURCE      Nicotiana tabacum (common tobacco)
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1705)
  AUTHORS   Yun,S.J. and Oh,S.H.
  TITLE     Cloning and characterization of a tobacco cDNA encoding
            calcium/calmodulin-dependent glutamate decarboxylase
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1705)
  AUTHORS   Yun,S.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-AUG-1997) Faculty of Biological Resources Science,
            Chonbuk National University, 664-14 1-ga Tokjin-dong, Chonju,
            Chonbuk 561-756, South Korea
FEATURES             Location/Qualifiers
     source          1..1705
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4097"
     gene            1..1705
                     /gene="NtGAD1"
     CDS             71..1561
                     /gene="NtGAD1"
                     /function="calmodulin binding protein"
                     /note="calcium-calmodulin-dependent enzyme"
                     /codon_start=1
                     /product="glutamate decarboxylase isozyme 1"
                     /protein_id="AAC24195.1"
                     /db_xref="GI:3252856"
                     /translation="MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECNKLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLGDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLIEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWAIWRNKEDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGFEGYKNVMENCQENARVLREGLEKSGRFNIISKEIGVPLVAFSLKDNSQHNEFEIS
                     ETLRRFGWIIPAYTMPPNAQHVTVLRVVIREDFSRTLAERLVIDIEKVLHELDTLPAR
                     VNAKLAVAEANGSGVHKKTDREVQLEITTAWKKFVADKKKKTNGVC"
BASE COUNT      523 a    301 c    385 g    496 t
ORIGIN

  Query Match            58.5%;  Score 883.4;  DB 8;  Length 1705;
  Best Local Similarity  75.2%;  Pred. No. 7.9e-247;
  Matches 1125;  Conservative    0;  Mismatches  351;  Indels   21;  Gaps    1;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||| || ||| |  | |  |||||    ||||||||| ||||||||||| ||||| ||
Db         71 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 130

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || || || |||||||| |||||| ||||| |||||||| || || ||||| || ||||||
Db        131 CGATATGTTCGTACTTCTCTTCCGAGGTTTAAGATGCCAGAGAATTCGATACCAAAGGAA 190

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              || || ||||| ||||| || || ||||| ||| | || || |||||| |  | || |||
Db        191 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGACTAAATTTA 250

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              || || |||||||| ||||||||||| || |||||| | ||||| || |||   |||||
Db        251 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAACAAACTGATGATGGATTCCATT 310
```

```
y        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
b        311 AACAAGAATTACGTTGACATGGATGAATACCCTGTAACCACTGAACTTCAGAATCGATGT 370

y        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
b        371 GTAAACATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 430

y        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
b        431 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 490

y        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
b        491 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTGTGACAAGCCCAATATTGTCACTGGT 550

y        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
b        551 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 610

y        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
b        611 GTAAAGTTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 670

y        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
b        671 GAGAACACAATTTGTGTAGCTGCTATCTTGGGTTCCACACTCAATGGTGAATTTGAAGAT 730

y        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
b        731 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 790

y        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
b        791 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 850

y        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
b        851 TGGGACTTTAGATTGCCATTGGTGAAGAGTATAAACGTGAGTGGTCACAAATATGGTCTT 910

y        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
b        911 GTTTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTACCTGACGAACTT 970

y        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
b        971 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1030

y        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
b       1031 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1090

y       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
b       1091 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1150

y       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
b       1151 AGTGGAAGATTCAACATAATATCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1210

y       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
b       1211 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1270

y       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
b       1271 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1330

y       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
b       1331 ATTAGAGAAGATTTCTCCCGTACACTCGCCGAGCGACTGGTAATAGACATTGAAAAAGTC 1390

y       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Db      1391 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTGGCCGAGGCG 1450

Qy      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
Db      1451 AATGG---------------------CAGCGGCGTGCATAAGAAAACAGATAGAGAAGTG 1489

Qy      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
Db      1490 CAGCTTGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACT 1546


RESULT 8
PETGADX
LOCUS       PETGADX                 1785 bp    mRNA    linear   PLN 15-OCT-1993
DEFINITION  Petunia hybrida glutamate decarboxylase (gad) mRNA, complete cds.
ACCESSION   L16797
VERSION     L16797.1  GI:294111
KEYWORDS    glutamate decarboxylase.
SOURCE      Petunia x hybrida
  ORGANISM  Petunia x hybrida
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Petunia.
REFERENCE   1  (bases 1 to 1785)
  AUTHORS   Baum,G., Chen,Y., Arazi,T., Takatsuji,H. and Fromm,H.
  TITLE     A plant glutamate decarboxylase containing a calmodulin binding
            domain. Cloning, sequence, and functional analysis
  JOURNAL   J. Biol. Chem. 268 (26), 19610-19617 (1993)
  MEDLINE   93374956
   PUBMED   8366104
COMMENT     Original source text: Petunia hybrida flower corolla cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1..1785
                     /organism="Petunia x hybrida"
                     /mol_type="mRNA"
                     /db_xref="taxon:4102"
                     /tissue_type="corolla"
                     /dev_stage="flower"
     gene            1..1785
                     /gene="gad"
     CDS             72..1574
                     /gene="gad"
                     /EC_number="4.1.1.15"
                     /codon_start=1
                     /product="glutamate decarboxylase"
                     /protein_id="AAA33709.1"
                     /db_xref="GI:294112"
                     /translation="MVLSKTVSQSDVSIHSTFASRYVRTSLPRFKMPDNSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLEDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLVEKNKETGWDTPIHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWVVWRNKDDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGYEGYKNVMENCQENASVLREGLEKTGRFNIISKEIGVPLVAFSLKDNRQHNEFEIS
                     ETLRRFGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVRDIEKVLHELDTLPAR
                     VNAKLAVAEEQAAANGSEVHKKTDSEVQLEMITAWKKFVEEKKKKTNRVC"
BASE COUNT      550 a    298 c    401 g    536 t
ORIGIN

  Query Match             58.4%;  Score 881.8;  DB 8;  Length 1785;
  Best Local Similarity   74.9%;  Pred. No. 2.3e-246;
  Matches 1121;  Conservative    0;  Mismatches  367;  Indels    9;  Gaps    1;

Qy         1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
Db        72 ATGGTTCTATCAAAGACAGTGTCGCAGAGCGATGTGTCCATTCACTCCACGTTTGCTTCT 131

Qy        61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
Db       132 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGATAATTCGATACCAAAAGAA 191

Qy       121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
```

```
           || || ||||||||||| || || || |||||| | || || || ||| || |||||||
b      192 GCAGCATATCAGATCATAAATGATGAACTGATGTTAGATGGAAACCCAAGGCTGAACTTG 251

y      181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
           || || ||||| || ||||||||||| || |||||||||||  | || |||  ||| ||
b      252 GCTTCTTTTGTTACAACATGGATGGAACCAGAGTGTGATAAGTTGATGATGGACTCTATT 311

y      241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
           ||||||||||||||||| ||||| || || || || ||||| || |||||||| ||||||
b      312 AACAAGAACTATGTTGATATGGATGAATATCCTGTTACCACTGAGCTTCAGAATCGATGT 371

y      301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
           || ||||||||| || ||| | || ||||||||  | ||||| |  || || || || |||
b      372 GTAAACATGATAGCTCATTTGTTTAATGCACCACTTGAAGATGGAGAAACTGCAGTTGGA 431

y      361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
           || ||||| |||||||| || || ||||| ||| | || || || || |||||| | |||
b      432 GTTGGAACTGTTGGATCCTCTGAAGCCATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 491

y      421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
           |||||||||||    |||||  |||||||||||   || || ||||||||||| || ||
       492 TGGCAGAACAAAATGAAAGCCCAAGGCAAACCCTGTGACAAGCCCAACATTGTTACTGGT 551

y      481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
           || ||||| || ||||| ||||||||||| || ||||| ||||| || || || ||||||
b      552 GCAAATGTCCAGGTGTGCTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 611

y      541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
           || ||  | ||||||||||||||||||||||||||| |  ||||||| || ||||| |||
b      612 GTAAAGCTTAGTGAAGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAGATGGTGGAT 671

y      601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
           || |||||||||||||| || |  ||  | |||||||| || ||||||||||| |||||
b      672 GAAAACACCATTTGTGTAGCTGCTATCTTAGGTTCCACCCTCAATGGAGAATTTGAAGAC 731

y      661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
           ||||| | |||||| |||||||||||||| ||||||||||||||||| ||||| || |||
       732 GTTAAGCGCTTGAATGATCTCTTGGTCGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 791

       721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
           || || |||||||| |||||||| ||||| |||||||||||  | || || ||  | ||
       792 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCGTTCATTTACCCAGAGCTTGAG 851

       781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
           |||||||||||| | || || |||||||| || ||||| ||||||||||| ||||| |||
       852 TGGGACTTTAGATTGCCATTAGTGAAGAGCATTAATGTAAGTGGTCACAAATATGGTCTT 911

       841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
           || || || || |||||||||||  | ||||| ||||| || ||||||||||| |||||
       912 GTCTATGCTGGTATTGGTTGGGTCGTTTGGAGGAACAAGGATGATTTGCCTGATGAACTT 971

       901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
           |||||||| || ||||||||||||||||| ||||| || || |||||||| || || |||
       972 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTTTCTAAA 1031

       961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
           ||||| || ||||| ||||||||||| ||||||||||| || |||||  | |||||||||
      1032 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTATGAGGGTTAC 1091

      1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
           |  |||||||||||||||||   ||| |||      || ||||| ||||| || || |||
      1092 AAGAATGTGATGGAGAATTGTCAAGAAAATGCATCGGTACTAAGAGAAGGGCTAGAAAAG 1151

      1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
           |||| ||| ||||||||  |||| || ||    ||||| ||  | || || |||||  |
      1152 ACAGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTACCTTTAGTAGCATTCTCTCTT 1211

      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
           ||||| | |||    ||||  |||||||| || || || |   |  | || | ||| |||
      1212 AAAGACAACAGGCAACACAACGAGTTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1271

      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
           || || || || || || ||||| ||||| ||||||||||| || |||||  | || ||

Db      1272 ATTGTTCCTGCATATACTATGCCACCAAACGCACAACACATTACAGTTCTCAGAGTTGTG 1331

Qy      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
             |||||||||||||||||  | || || || ||  |||| || |  || || || |||||
Db      1332 ATCAGAGAAGATTTCTCCCGTACGCTTGCAGAACGACTGGTAAGAGACATCGAAAAAGTC 1391

Qy      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
              | | ||| || ||    || ||| |  | || | |   ||  |  |  | |   | |||
Db      1392 CTTCATGAACTTGACACACTCCCTGCACGTGTCAATGCTAAGCTCGCTGTGGCCGAGGAG 1451

Qy      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
              ||   |   | ||  ||         |  |    |   ||||| |  |||| |||   |
Db      1452 CAGGCGGCTGCGAATGGC---------AGCGAGGTGCATAAGAAAACAGATAGCGAAGTG 1502

Qy      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
             |||   || || || ||||  |||||||||||||| |  || | |||||||| || |
Db      1503 CAGTTGGAGATGATAACTGCATGGAAGAAGTTTGTTGAAGAAAAGAAGAAGAAGACT 1559


RESULT 9
PETGLUDECA
LOCUS       PETGLUDECA              1785 bp    mRNA    linear   PLN 26-JUL-1993
DEFINITION  Petunia hybrida glutamate decarboxylase (gad) mRNA, complete cds.
ACCESSION   L16977
VERSION     L16977.1  GI:309679
KEYWORDS    calmodulin binding protein; glutamate decarboxylase.
SOURCE      Petunia x hybrida
  ORGANISM  Petunia x hybrida
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Petunia.
REFERENCE   1  (bases 1 to 1785)
  AUTHORS   Baum,G., Chen,Y., Arazi,T., Takatsuji,H. and Fromm,H.
  TITLE     A plant glutamate decarboxylase containing a calmodulin-binding
            domain: cloning sequence and functional analysis
  JOURNAL   J. Biol. Chem. (1993) In press
COMMENT     Original source text: Petunia hybrida flower corolla cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1. .1785
                     /organism="Petunia x hybrida"
                     /mol_type="mRNA"
                     /db_xref="taxon:4102"
                     /tissue_type="corolla"
                     /dev_stage="flower"
     gene            1. .1785
                     /gene="gad"
     CDS             72. .1574
                     /gene="gad"
                     /EC_number="4.1.1.15"
                     /codon_start=1
                     /product="glutamate decarboxylase"
                     /protein_id="AAA33710.1"
                     /db_xref="GI:309680"
                     /translation="MVLSKTVSQSDVSIHSTFASRYVRTSLPRFKMPDNSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLEDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLVEKNKETGWDTPIHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWVVWRNKDDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGYEGYKNVMENCQENASVLREGLEKTGRFNIISKEIGVPLVAFSLKDNRQHNEFEIS
                     ETLRRFGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVRDIEKVLHELDTLPAR
                     VNAKLAVAEEQAAANGSEVHKKTDSEVQLEMITAWKKFVEEKKKKTNRVC"
BASE COUNT      550 a    298 c    401 g    536 t
ORIGIN

  Query Match             58.4%;  Score 881.8;  DB 8;  Length 1785;
  Best Local Similarity   74.9%;  Pred. No. 2.3e-246;
  Matches 1121;  Conservative    0;  Mismatches  367;  Indels    9;  Gaps    1;

Qy         1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
             ||||| || ||  |  | || ||| ||   || || ||| | |||||||| || || ||
Db        72 ATGGTTCTATCAAAGACAGTGTCGCAGAGCGATGTGTCCATTCACTCCACGTTTGCTTCT 131
```

```
y          61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || || || || ||||| ||||| ||||| || ||||| || || ||||| || || |||
          132 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGATAATTCGATACCAAAAGAA 191

y         121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              || || ||||||||||| || || || ||||| | || || || ||| || |||||||
b         192 GCAGCATATCAGATCATAAATGATGAACTGATGTTAGATGGAAACCCAAGGCTGAACTTG 251

y         181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              || || ||||| || ||||||||||| || |||||||||||  | || |||  ||| ||
          252 GCTTCTTTTGTTACAACATGGATGGAACCAGAGTGTGATAAGTTGATGATGGACTCTATT 311

y         241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||||||||| ||||| || || || || ||||| || |||||||| ||||||
          312 AACAAGAACTATGTTGATATGGATGAATATCCTGTTACCACTGAGCTTCAGAATCGATGT 371

              301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              || ||||||||| || ||| | || ||||||||  | ||||| |  || || || || |||
          372 GTAAACATGATAGCTCATTTGTTTAATGCACCACTTGAAGATGGAGAAACTGCAGTTGGA 431

          361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              || ||||| |||||||| || || ||||| ||| | || || || || |||||| | |||
          432 GTTGGAACTGTTGGATCCTCTGAAGCCATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 491

          421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              |||||||||||    |||||  |||||||||||   || || ||||||||||| || ||
          492 TGGCAGAACAAAATGAAAGCCCAAGGCAAACCCTGTGACAAGCCCAACATTGTTACTGGT 551

          481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              || ||||| || ||||| ||||||||||| || ||||| ||||| || || || ||||||
          552 GCAAATGTCCAGGTGTGCTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 611

          541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              || ||  | ||||||||||||||||||||||||||| |  ||||||| || ||||| |||
          612 GTAAAGCTTAGTGAAGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAGATGGTGGAT 671

          601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              || |||||||||||||| || |  ||  | |||||||| || ||||||||||| |||||
          672 GAAAACACCATTTGTGTAGCTGCTATCTTAGGTTCCACCCTCAATGGAGAATTTGAAGAC 731

          661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| | |||||| ||||||||||||||||| ||||||||||||||||| ||||| || |||
          732 GTTAAGCGCTTGAATGATCTCTTGGTCGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 791

          721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || || |||||||| |||||||| ||||| |||||||||||  | || ||  || | ||
          792 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCGTTCATTTACCCAGAGCTTGAG 851

          781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              |||||||||||| | || || |||||||| || ||||| ||||||||||| ||||| |||
          852 TGGGACTTTAGATTGCCATTAGTGAAGAGCATTAATGTAAGTGGTCACAAATATGGTCTT 911

          841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || || || || |||||||||||  | ||||| ||||| || ||||||||||| |||||
          912 GTCTATGCTGGTATTGGTTGGGTCGTTTGGAGGAACAAGGATGATTTGCCTGATGAACTT 971

          901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              |||||||| || ||||||||||||||||| ||||| || || ||||||||| || || |||
          972 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTTTCTAAA 1031

          961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              ||||| || ||||| ||||||||||| ||||||||||| || |||||  | |||||||||
         1032 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTATGAGGGTTAC 1091

         1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              |  |||||||||||||||||   ||| |||      || ||||| ||||| || || |||
         1092 AAGAATGTGATGGAGAATTGTCAAGAAAATGCATCGGTACTAAGAGAAGGGCTAGAAAAG 1151

         1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              |||| ||| ||||||||  |||| || ||    ||||| ||  | || || |||||  |
         1152 ACAGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTACCTTTAGTAGCATTCTCTCTT 1211

         1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              ||||| | |||    |||| |||||||| || || || |   |  | || | ||| |||
Db       1212 AAAGACAACAGGCAACACAACGAGTTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1271

Qy       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              || || || || || || ||||| ||||| ||||||||||| || |||||  | || ||
Db       1272 ATTGTTCCTGCATATACTATGCCACCAAACGCACAACACATTACAGTTCTCAGAGTTGTG 1331

Qy       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              |||||||||||||||||  | || || || ||  |||| || |  || || || |||||
Db       1332 ATCAGAGAAGATTTCTCCCGTACGCTTGCAGAACGACTGGTAAGAGACATCGAAAAAGTC 1391

Qy       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
               | | ||| || ||    || ||| |  | || | |   ||  |  |  |  |   | |||
Db       1392 CTTCATGAACTTGACACACTCCCTGCACGTGTCAATGCTAAGCTCGCTGTGGCCGAGGAG 1451

Qy       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
               ||   |   | ||  ||         |  |    |   ||||| |  |||| |||   |
Db       1452 CAGGCGGCTGCGAATGGC---------AGCGAGGTGCATAAGAAAACAGATAGCGAAGTG 1502

Qy       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
              |||   || || || ||||  |||||||||||||| |  || | |||||||| || |
Db       1503 CAGTTGGAGATGATAACTGCATGGAAGAAGTTTGTTGAAGAAAAGAAGAAGAAGACT 1559


RESULT 10
NTU54774
LOCUS       NTU54774                1929 bp    mRNA    linear   PLN 14-JAN-1997
DEFINITION  Nicotiana tabacum glutamate decarboxylase (NtGAD1) mRNA, complete
            cds.
ACCESSION   U54774
VERSION     U54774.1  GI:1777920
KEYWORDS    .
SOURCE      Nicotiana tabacum (common tobacco)
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1929)
  AUTHORS   Dharmasiri,M.A.N., Lu,Y.T. and Harrington,H.M.
  TITLE     Cloning and sequencing of a tobacco cDNA encoding glutamate
            decarboxylase
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1929)
  AUTHORS   Dharmasiri,M.A.N., Lu,Y.T. and Harrington,H.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-APR-1996) M.A. Nihal Dharmasiri, Plant Molecular
            Physiology, University of Hawaii, 3190 Maile Way, Honolulu, HI
            96822, USA
FEATURES             Location/Qualifiers
     source          1..1929
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /strain="Xanthi"
                     /db_xref="taxon:4097"
     gene            1..1929
                     /gene="NtGAD1"
     CDS             66..1556
                     /gene="NtGAD1"
                     /note="calmodulin regulated enzyme; calmodulin-binding
                     protein"
                     /codon_start=1
                     /product="glutamate decarboxylase"
                     /protein_id="AAB40608.1"
                     /db_xref="GI:1777921"
                     /translation="MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECNKLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLGDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLIEKNKETGWDTPIHVDAASGEFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWAIWRNKEDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGFEGYKNVMENCQENARVLREGLEKSGRFNIISKEIGVPLVAFSLKDNSQHNEFEIS
                     ETLRRFGWIIPAYTMPPNAQHVTVLRVVIREDFSRTLAERLVIDIEKVLHELDTLPAR
                     VNAKLAVAEANGSGVHKKTDREVQLEITTAWKKFVADKKKKTNGVC"
```

```
SE COUNT      607 a    323 c    416 g    583 t
IGIN

Query Match            58.4%;  Score 881.8;  DB 8;  Length 1929;
Best Local Similarity  75.1%;  Pred. No. 2.4e-246;
Matches 1124;  Conservative    0;  Mismatches  352;  Indels   21;  Gaps    1;

   1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
     ||||| || ||| |  | |  |||||    ||||||||| |||||||||| ||||| ||
  66 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 125

  61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
     || || || |||||||| ||||| ||||| |||||||| || || ||||| || ||||||
 126 CGATATGTTCGTACTTCTCTTCCGAGGTTTAAGATGCCAGAGAATTCGATACCAAAGGAA 185

 121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
     || || ||||| ||||| || || ||||| ||| | || || |||||| |  | || |||
 186 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGACTAAATTTA 245

 181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
     || || |||||||| ||||||||||| || |||||| | ||||| || |||   |||||
 246 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAACAAACTGATGATGGATTCCATT 305

 241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
     |||||||| || ||||||||||| || ||||| || ||||| ||||||||||| ||||||
 306 AACAAGAATTACGTTGACATGGATGAATACCCTGTAACCACTGAACTTCAGAATCGATGT 365

 301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
     || |||||||| || ||| | || || |||||  | | ||| |  ||||| || || |||
 366 GTAAACATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 425

 361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
     || ||||| |||||||| || ||||| || ||| | || || || || |||||| | |||
 426 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 485

 421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
     ||||| || ||    |||||  ||||||| |||   || || ||||| |||||||| ||
 486 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTGTGACAAGCCCAATATTGTCACTGGT 545

 481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
     |||||||| || ||||||||||||||||| || ||||| ||||| || || || ||||||
 546 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 605

 541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
     || || |||||||| ||||||||||||||||||||| |  ||||||| || ||||| |||
 606 GTAAAGTTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 665

 601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
     |||||||| |||||||| || |  ||  | |||||||| || ||||| ||||| ||||||
 666 GAGAACACAATTTGTGTAGCTGCTATCTTGGGTTCCACACTCAATGGTGAATTTGAAGAT 725

 661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
     ||||| | |||||| || |||||| | || ||||||||||||||||| ||||| || |||
 726 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 785

 721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
     || || |||||||| |||||||| | ||| |||||||| ||  | ||||| ||  | |||
 786 ATTCATGTGGATGCAGCAAGTGGTGAATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 845

 781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
     |||||||||||| | || |||||||||||||| || |||||||||||||| ||||| |||
 846 TGGGACTTTAGATTGCCATTGGTGAAGAGTATTAACGTGAGTGGTCACAAATATGGTCTT 905

 841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
     || || || || ||||||||||  || ||||| || || || || || ||||| |||||
 906 GTTTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTACCTGACGAACTT 965

 901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
     |||||||| || ||||||||||||||||| ||||| || || |||||||| ||||| |||
 966 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1025

 961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
     ||||| || ||||| ||||||||||| ||||||||||| || |||||    |||||||||
```

```
Db        1026 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1085

Qy        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
               |  ||||| |||||||||||   ||| |||   |  || ||||| ||||||||||| ||
Db        1086 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1145

Qy        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
               |  | ||| ||||||||  | || || ||    ||||| ||| | || ||||||||  |
Db        1146 AGTGGAAGATTCAACATAATATCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1205

Qy        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
               ||||| | |||    |||| |||||||||||| || || |  ||| | || | |||||||
Db        1206 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1265

Qy        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
               ||  | || || || || ||||| |||||||| |||||  |||| |||||  | || ||
Db        1266 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1325

Qy        1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
               || ||||||||||||||  | |||||||| ||| |||| || || || || || |||||
Db        1326 ATTAGAGAAGATTTCTCCCGTACACTCGCCGAGCGACTGGTAATAGACATTGAAAAAGTC 1385

Qy        1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
                | |  ||||| ||    |||||  |||| || |     ||  || |  | |   | | |
Db        1386 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTGGCCGAGGCG 1445

Qy        1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
               ||  |                      |  | |  |   ||||| |  ||||  ||   |
Db        1446 AATGG---------------------CAGCGGCGTGCATAAGAAAACAGATAGAGAAGTG 1484

Qy        1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
               |||   || || |  ||||  |||||||| ||||| || || | |||||||| || |
Db        1485 CAGCTTGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACT 1541


RESULT 11
AF352732
LOCUS       AF352732                1672 bp    mRNA    linear   PLN 02-MAY-2003
DEFINITION  Nicotiana tabacum glutamate decarboxylase isozyme 1 mRNA, complete
            cds.
ACCESSION   AF352732
VERSION     AF352732.1  GI:13310812
KEYWORDS    .
SOURCE      Nicotiana tabacum (common tobacco)
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1672)
  AUTHORS   Yevtushenko,D., McLean,M.D., Peiris,S.E., Van Cauwenberghe,O.R. and
            Shelp,B.J.
  TITLE     Two isoforms of tobacco glutamate decarboxylase are regulated by
            calcium/calmodulin and differ in organ distribution
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1672)
  AUTHORS   Yevtushenko,D., McLean,M.D. and Shelp,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-FEB-2001) Department of Plant Agriculture, University
            of Guelph, Division of Biotechnology, Guelph, Ontario N1G 2W1,
            Canada
FEATURES             Location/Qualifiers
     source          1..1672
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /cultivar="Samsun NN"
                     /db_xref="taxon:4097"
     CDS             4..1494
                     /EC_number="4.1.1.15"
                     /function="converts glutamate to gamma-aminobutyric acid"
                     /note="GAD; GAD3; NtGAD3; calcium/calmodulin-dependent
                     enzyme"
                     /codon_start=1
                     /product="glutamate decarboxylase isozyme 1"
```

```
                     /protein_id="AAK18620.1"
                     /db_xref="GI:13310813"
                     /translation="MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECNKLMMDSINKNYVDMGEYPVTTELQNRCVN
                     MIAHLFNAPLGDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLIEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWAIWRNKEDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGFEGYKNVMENCQENARVLREGLEKSGRFNIISKEIGVPLVAFSLKDNSQHNEFEIS
                     ETLRRFGWIIPAYTMPPNAQHVTVLRVVIREDFSRTLAERLVIDIEKVLHELDTLPAR
                     VNAKLAVAEANGSGVHKKTDREVQLEITAAWKKFVADKKKKTNGVC"
SE COUNT      511 a    280 c    392 g    483 t      6 others
IGIN

Query Match             58.1%;  Score 877;  DB 8;  Length 1672;
Best Local Similarity   74.9%;  Pred. No. 5.9e-245;
Matches 1121;  Conservative    0;  Mismatches  355;  Indels   21;  Gaps    1;

   1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
     ||||| || ||| |  | |  |||||    ||||||||| |||||||||| ||||| ||
   4 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 63

  61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
     || || || || ||||| ||||| ||||| |||||||| || || || || || || |||
  64 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAGATGCCAGAGAATTCTATACCAAAAGAA 123

 121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
     || || ||||| ||||| || || ||||| ||| | || || |||||| || | || |||
 124 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 183

 181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
     || || |||||||| ||||||||||| || |||||| ||||  | || |||   |||||
 184 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAATAAGTTAATGATGGATTCCATT 243

 241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
     ||||||||||| ||||||||||  || ||||| || ||||| || |||||||| ||||||
 244 AACAAGAACTACGTTGACATGGGTGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 303

 301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
     || || ||||| || ||| | || || |||||  | | ||| |  ||||| || || |||
 304 GTAAATATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 363

 361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
     || ||||| |||||||| || ||||| || ||| | || || || ||||||||| | |||
 364 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCCTTCAAGAGAAAA 423

 421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
     ||||| || ||    |||||  |||| || |||  ||| || ||||| |||||||| ||
 424 TGGCAAAATAAAATGAAAGCCCAAGGGAAGCCCTGCGACAAGCCCAATATTGTCACTGGT 483

 481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
     |||||||| || ||||||||||||||||| || ||||| ||||| || ||  | || |||
 484 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 543

 541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
     || ||||||||||| ||||||||||||||||||||| |  ||||||| || ||||| |||
 544 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 603

 601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
     ||||| || |||||||| || |  ||  | ||||||||||| ||||| ||||| ||||||
 604 GAGAATACAATTTGTGTAGCTGCTATCTTGGGTTCCACTCTCAATGGTGAATTTGAAGAT 663

 661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
     ||||| | |||||| || |||||| | || ||||||||||||||||| ||||| || |||
 664 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 723

 721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
     || || |||||||| |||||||| |||||||||||||| ||  | ||||| ||  | |||
 724 ATTCATGTGGATGCAGCAAGTGGTGGATTCATTGCACCATTCCTTTATCCTGAGCTTGAA 783

 781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
     ||||| |||||| | || |||||||||||||| ||||||||||||||||| ||||| |||
 784 TGGGATTTTAGATTACCATTGGTGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 843
```

```
Qy       841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
             || || || || ||||||||||  || ||||| || || || || |||||||| |||||
Db       844 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 903

Qy       901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
             || ||||| || ||||||||||||||||| ||||| || || |||||||| ||||| |||
Db       904 ATTTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 963

Qy       961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
             ||||| || ||||| ||||||||||| ||||||||||| || |||||    |||||||||
Db       964 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1023

Qy      1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
             |  ||||| |||||||||||   ||| |||   |  || ||||| ||||||||||| ||
Db      1024 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1083

Qy      1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
             |  | ||| ||||| ||  |||| || ||    ||||| ||| | || ||||||||  |
Db      1084 AGTGGAAGATTCAATATAATCTCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1143

Qy      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
             ||||| | |||    |||| |||||||||||| || || |  ||| | || | |||||||
Db      1144 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1203

Qy      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
             ||  | || || || || ||||| |||||||| |||||  |||| |||||  | || |||
Db      1204 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTT 1263

Qy      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
             || ||||||||||||||  | |||||||| ||| |||| ||||| || || || |||||
Db      1264 ATTAGAGAAGATTTCTCCCGTACACTCGCGGAGCGACTGGTGATAGACATTGAAAAAGTC 1323

Qy      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
              | |  ||||| ||    |||||  |||| || |     ||  |  |  | |   | | |
Db      1324 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTCGCCGTGGCCGAGGCG 1383

Qy      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
             ||  |                      |  | |  |   ||||| |  ||||  ||   |
Db      1384 AATGG---------------------CAGCGGCGTGCATAAGAAAACAGATAGAGAAGTG 1422

Qy      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
             |||   || || |   |||  |||||||| ||||| || || | |||||||| || |
Db      1423 CAGCTGGAGATTACTGCTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACT 1479
```

```
RESULT 12
AY124873
LOCUS       AY124873                1485 bp    mRNA    linear   PLN 06-JUL-2002
DEFINITION  Arabidopsis thaliana At1g65960/F12P19_12 mRNA, complete cds.
ACCESSION   AY124873
VERSION     AY124873.1  GI:21700916
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1485)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L.,
            Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
            Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1485)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L.,
            Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
```

```
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
            Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
TITLE       Direct Submission
JOURNAL     Submitted (17-JUN-2002) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
MMENT       RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
            Chen,H., Kim,C.J., Shinn,P., Banh,J. Bowser,L., Chan,M.M.,
            Chang,E., Dale,J.M., Deng, J.M., Goldsmith,A.D., Jones,T.,
            Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamada,K., Yamamura, Y., Yu,G., Davis,R.W.,
            Theologis,A., and Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
ATURES               Location/Qualifiers
   source            1..1485
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="U14793"
                     /note="This clone is in pUNI 51
                     ecotype: Columbia"
   CDS               1..1485
                     /note="glutamate decarboxylase 2"
                     /codon_start=1
                     /product="At1g65960/F12P19_12"
                     /protein_id="AAM70582.1"
                     /db_xref="GI:21700917"
                     /translation="MVLTKTATNDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQI
                     IKDELMLDGNPRLNLASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNI
                     IARLFNAPLEESETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGA
                     NVQVCWEKFARYFEVELKEVNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFE
                     DVKRLNDLLVKKNEETGWNTPIHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHK
                     YGLVYAGIGWVVWRAAEDLPEELIFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRL
                     GFEGYKNVMENCIENMVVLKEGIEKTERFNIVSKDQGVPVVAFSLKDHSFHNEFEISE
                     MLRRFGWIVPAYTMPADAQHITVLRVVIREDFSRTLAERLVADISKVLHELDTLPSKI
                     SKKMGIEGIAENVKEKKMEKEILMEVIVGWRKFVKERKKMNGVC"
SE COUNT       438 a    290 c    385 g    372 t
IGIN

Query Match             57.7%;  Score 870.2;  DB 8;  Length 1485;
Best Local Similarity   75.1%;  Pred. No. 5.6e-243;
Matches 1110;  Conservative    0;  Mismatches  348;  Indels   21;  Gaps    1;

     31 GACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCACTTCCTAGGTTC 90
        || |  || |||  | |||  |||| ||| || || || || ||| ||||||| | ||
     28 GATGAATCTGTCTGCACCATGTTCGGATCTCGCTATGTTCGCACTACACTTCCCAAGTAT 87

     91 AAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTG 150
         ||||    || || ||||| || || || || |||||||||||||| || || ||||||
     88 GAGATTGGTGAGAATTCGATACCGAAAGACGCTGCGTATCAGATCATAAAAGATGAGCTG 147

    151 ATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCT 210
        |||||||| || || ||| || | ||| |||| || |||||||| ||||||||||| ||
    148 ATGCTTGATGGTAACCCAAGGCTTAACCTAGCTTCTTTTGTGACTACATGGATGGAACCA 207

    211 GAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTAC 270
        |||||||| ||||||||||||  ||| ||||| |||||||| ||||| ||||| ||||||
```

```
Db    208 GAGTGTGACAAACTCATCATGGACTCTATCAATAAGAACTACGTTGATATGGATGAGTAC 267

Qy    271 CCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATCTATTCAATGCA 330
          || ||||| || || || |||||||||||||| ||||| || || |  || ||||||||
Db    268 CCTGTCACAACTGAGCTCCAGAACCGATGTGTAAACATTATAGCTCGACTGTTCAATGCG 327

Qy    331 CCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATA 390
          ||  | || ||  | ||||| || || |||||||| || ||||| || || || |||||
Db    328 CCACTCGAGGAATCTGAGACGGCGGTGGGAGTAGGGACAGTTGGTTCTTCAGAAGCCATC 387

Qy    391 ATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAA 450
          ||||| ||||| |||||||||||  | |||||||||||||| ||||| ||||| || |||
Db    388 ATGTTAGCCGGATTGGCCTTCAAAAGAAAATGGCAGAACAAACGCAAGGCTGAGGGTAAA 447

Qy    451 CCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTC 510
          |||   || ||||||||||||||||||||||||||||||||||| || ||||||||||||
Db    448 CCCTATGACAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTTTGCTGGGAGAAATTC 507

Qy    511 GCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATG 570
          ||| ||||||| ||||| || || |||||||| ||  | |||||||| ||||| ||||||
Db    508 GCTCGGTACTTCGAGGTGGAGCTAAAGGAAGTAAACCTAAGTGAAGGTTACTACGTGATG 567

Qy    571 GACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTT 630
          || ||  |  |||| |  || ||||| || |||||||| || ||||| || | |||  |
Db    568 GATCCAGACAAAGCAGCAGAAATGGTAGACGAGAACACAATCTGTGTCGCAGCCATATTG 627

Qy    631 GGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAA 690
          || ||||| || || || || |||||||| || ||||   | || ||  |  | ||| |
Db    628 GGATCCACACTCAACGGTGAGTTCGAAGACGTGAAACGTCTCAATGACTTGCTAGTCAAG 687

Qy    691 AAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTC 750
          || ||| | || || || ||| | ||||| |||||||||||||| ||||||||||| |||
Db    688 AAAAACGAGGAGACTGGTTGGAACACACCGATCCACGTGGATGCAGCAAGTGGAGGGTTC 747

Qy    751 ATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGT 810
          || || |||||| | ||||| ||||| |||||||||||||| ||||| ||||| |||||
Db    748 ATAGCTCCGTTTATCTATCCTGAATTAGAATGGGACTTTAGGCTTCCTTTGGTTAAGAGC 807

Qy    811 ATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGG 870
          ||||| ||||||||||||||||||||||| || || || || |||||||||||  | |||
Db    808 ATCAACGTGAGTGGTCACAAGTATGGACTAGTCTATGCTGGTATTGGTTGGGTCGTGTGG 867

Qy    871 AGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGAC 930
          ||      ||||||||| ||||| || || ||||| ||||| |||||||||||||||||
Db    868 AGGGCAGCAGAGGATTTACCTGAAGAGCTTATCTTTCATATTAATTATCTTGGTGCTGAT 927

Qy    931 CAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTAC 990
          |||||||| || ||||||||||||||||| || || || ||| | |||||||||||||||
Db    928 CAACCCACTTTCACTCTCAATTTCTCCAAGGGATCGAGCCAAATTATTGCTCAATACTAC 987

Qy    991 CAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAAT 1050
          || || || ||  | ||   |||||| |||| |||||||||||||||||||| ||||||
Db    988 CAGCTCATTCGTCTTGGATTCGAGGGGTACAAAAATGTGATGGAGAATTGCATAGAGAAC 1047

Qy   1051 ATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGAC 1110
          ||| | || || |  |||||  | ||||| |||||  | |||||||| ||||||||||||
Db   1048 ATGGTGGTTCTCAAAGAAGGTATAGAGAAAACAGAGCGTTTCAACATAGTCTCAAAGGAC 1107

Qy   1111 GAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAA 1170
           | ||||||||| | ||||| |||||  | || ||    || |  ||||  ||||||||
Db   1108 CAAGGAGTGCCAGTCGTCGCCTTCTCTCTCAAGGACCATAGTTTCCACAACGAGTTCGAG 1167

Qy   1171 ATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAAT 1230
          ||||| || ||||| ||  | | ||| ||||| || || || ||||| |||||| |  ||
Db   1168 ATCTCTGAGATGCTACGTCGTTTTGGCTGGATCGTCCCAGCTTACACTATGCCTGCCGAT 1227

Qy   1231 GCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCT 1290
          ||||| |||||||| ||||| ||||| || ||||| ||||||||||| |||||||||||
Db   1228 GCACAGCACATCACGGTTCTGCGTGTTGTCATCAGGGAAGATTTCTCAAGAACACTCGCG 1287

Qy   1291 GAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGCTTCCTTCGAGA 1350
          |||||||||||    |||||   ||| ||| | | |||||| |||    | ||||| |
Db   1288 GAGAGACTTGTTGCTGATATTTCGAAGGTGCTTCATGAGCTAGATACCTTGCCTTCCAAG 1347
```

```
       1351 GTGATTCACAAAATATCACTTGGACAAGAGAAGAGTGAATCTAACAGCGATAACTTGATG 1410
             |                     |    ||||||  | |    |  |          |
       1348 AT---------------------ATCTAAGAAGATGGGAATAGAAGGGATCGCGGAAAAT 1386

       1411 GTCACGGTGAAGAAGAGCGATATCGACAAGCAGAGAGATATCATCACTGGCTGGAAGAAG 1470
            || | || ||||||||  || |  || |  | ||  ||  | ||   ||| |||| ||||
       1387 GTAAAGGAGAAGAAGATGGAGAAGGAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAG 1446

       1471 TTTGTCGCCGACAGGAAGAAGACGAGTGGTATCTGCTAA 1509
            |||||    || |||||||||| || |||| | ||||||
       1447 TTTGTGAAGGAGAGGAAGAAGATGAATGGTGTGTGCTAA 1485


SULT 13
U46665
CUS         ATU46665               1676 bp    mRNA    linear   PLN 01-SEP-1998
FINITION    Arabidopsis thaliana glutamate decarboxylase 2 (GAD2) mRNA,
            complete cds.
CESSION     U46665
RSION       U46665.1  GI:1184959
YWORDS      .
URCE        Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
FERENCE     1  (bases 1 to 1676)
AUTHORS     Turano,F.J. and Fang,T.K.
TITLE       Characterization of two glutamate decarboxylase cDNA clones from
            Arabidopsis
JOURNAL     Plant Physiol. 117 (4), 1411-1421 (1998)
MEDLINE     98369120
 PUBMED     9701597
FERENCE     2  (bases 1 to 1676)
AUTHORS     Turano,F.J. and Thakkar,S.
TITLE       Direct Submission
JOURNAL     Submitted (21-JAN-1996) Frank J. Turano, Climate Stress Lab,
            USDA/ARS, BARC-West, Bldg. 046A, Beltsville, MD 20705, USA
ATURES               Location/Qualifiers
   source            1. .1676
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /strain="ecotype Columbia"
                     /db_xref="taxon:3702"
   gene              1. .1676
                     /gene="GAD2"
   CDS               17. .1501
                     /gene="GAD2"
                     /codon_start=1
                     /product="glutamate decarboxylase 2"
                     /protein_id="AAC33485.1"
                     /db_xref="GI:1184960"
                     /translation="MVLTKTATNDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQI
                     IKDELMLDGNPRLNLASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNI
                     IARLFNAPLEESETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGA
                     NVQVCWEKFARYFEVELKEVNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFE
                     DVKRLNDLLVKKNEETGWNTPIHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHK
                     YGLVYAGIGWVVWRAAEDLPEELIFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRL
                     GFEGYKNVMENCIENMVVLKEGIEKTERFNIVSKDQGVPVVAFSLKDHSFHNEFEISE
                     MLRRFGWIVPAYTMPADAQHITVLRVVIREDFSRTLAERLVADISKVLHELDTLPSKI
                     SKKMGIEGIAENVKEKKMEKEILMEVIVGWRKFVKERKKMNGVC"
ASE COUNT      498 a    313 c    424 g    441 t
IGIN

 Query Match             57.7%;  Score 870.2;  DB 8;  Length 1676;
 Best Local Similarity   75.1%;  Pred. No. 5.7e-243;
 Matches 1110;  Conservative    0;  Mismatches  348;  Indels   21;  Gaps    1;

y         31 GACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCACTTCCTAGGTTC 90
             || |  || |||  | |||  |||| |||||| || || || ||| |||||||||| | ||
b         44 GATGAATCTGTCTGCACCATGTTCGGATCTCGCTATGTTCGCACTACACTTCCCAAGTAT 103
```

```
Qy        91 AAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTG 150
              ||||    || || ||||| || || || || || |||||||||||| || || ||||||
Db       104 GAGATTGGTGAGAATTCGATACCGAAAGACGCTGCATATCAGATCATAAAAGATGAGCTG 163

Qy       151 ATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCT 210
             |||||||| || || ||  || | ||| |||| || |||||||| ||||||||||||| ||
Db       164 ATGCTTGATGGTAACCCGAGGCTTAACCTAGCTTCGTTTGTGACTACATGGATGGAACCA 223

Qy       211 GAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTAC 270
             |||||||| ||||||||||||  ||| |||||||||||||| ||||| |||||| ||||||
Db       224 GAGTGTGACAAACTCATCATGGACTCTATCAACAAGAACTACGTTGATATGGATGAGTAC 283

Qy       271 CCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATCTATTCAATGCA 330
             || ||||| || || || ||||||||||||||| ||||| || || |  || ||||||||
Db       284 CCTGTCACAACTGAGCTCCAGAACCGATGTGTAAACATTATAGCTCGACTGTTCAATGCG 343

Qy       331 CCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATA 390
             ||  | || ||  | ||||| || || |||||||| || |||| || ||  || |||||
Db       344 CCACTCGAGGAATCTGAGACGGCGGTGGGAGTAGGGACAGTTGGTTCTTCAGAAGCCATC 403

Qy       391 ATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAA 450
             ||||| ||||| ||||||||||||  | |||||||||||||| ||||| |||||| || |||
Db       404 ATGTTAGCCGGATTGGCCTTCAAAAGAAAATGGCAGAACAAACGCAAGGCTGAGGGTAAA 463

Qy       451 CCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTC 510
             |||   || |||||||||||||||||| |||||||||||||||| || ||||||||||||
Db       464 CCCTATGACAAACCCAACATTGTCACTGGAGCCAATGTTCAAGTTTGCTGGGAGAAATTC 523

Qy       511 GCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATG 570
             ||| ||||||| ||||| || || |||||||| ||  | |||||||| ||||| ||||||
Db       524 GCTCGGTACTTCGAGGTGGAGCTAAAGGAAGTAAACCTAAGTGAAGGTTACTACGTGATG 583

Qy       571 GACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTT 630
             || ||  |  |||| |  || ||||| || |||||||| || ||||| || | |||  |
Db       584 GATCCAGACAAAGCAGCAGAAATGGTAGACGAGAACACAATCTGTGTCGCAGCCATATTG 643

Qy       631 GGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAA 690
             || ||||| || || || || |||||||| || ||||   | || ||  |  | ||| |
Db       644 GGATCCACACTCAACGGTGAGTTCGAAGACGTGAAACGTCTCAATGACTTGCTAGTCAAG 703

Qy       691 AAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTC 750
             || ||| | || || || || | | ||||| |||||||||||||| |||||||||||| |||
Db       704 AAAAACGAGGAGACTGGTTGGAACACACCGATCCACGTGGATGCAGCAAGTGGAGGGTTC 763

Qy       751 ATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGT 810
             || || |||||| | ||||| ||||| |||||||||||||||||||||| ||||| ||||||
Db       764 ATAGCTCCGTTTATCTATCCTGAATTAGAATGGGACTTTAGACTTCCTTTGGTTAAGAGT 823

Qy       811 ATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGG 870
             ||||| ||||||||||||||||||||||||| || || || | ||||||||||||  | |||
Db       824 ATCAACGTGAGTGGTCACAAGTATGGACTGGTCTATGCTGGTATTGGTTGGGTCGTGTGG 883

Qy       871 AGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGAC 930
             ||      |||||||||||||| || || ||||| ||||| ||||||||||||||||||
Db       884 AGGGCAGCAGAGGATTTGCCTGAAGAGCTTATCTTTCATATTAATTATCTTGGTGCTGAT 943

Qy       931 CAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTAC 990
             |||||||| || |||||||||||||||||  || || || ||| | ||||||||||||||||
Db       944 CAACCCACTTTCACTCTCAATTTCTCCAAGGGATCGAGCCAAATTATTGCTCAATACTAC 1003

Qy       991 CAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAAT 1050
             || || || ||  | ||   |||||| |||| |||||||||||||||||||| ||||||
Db      1004 CAGCTCATTCGTCTTGGATTCGAGGGGTACAAAAATGTGATGGAGAATTGCATAGAGAAC 1063

Qy      1051 ATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGAC 1110
             ||| | || || |  |||||  | ||||| |||||  | |||||||| ||||||||||||
Db      1064 ATGGTGGTTCTCAAAGAAGGGATAGAGAAAACAGAGCGTTTCAACATAGTCTCAAAGGAC 1123

Qy      1111 GAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAA 1170
              | ||||||||| | || || |||||  | || ||    || |  ||||  ||||||||
Db      1124 CAAGGAGTGCCAGTCGTAGCCTTCTCTCTCAAGGACCATAGTTTCCACAACGAGTTCGAG 1183

Qy      1171 ATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAAT 1230
```

```
     ||||| || ||||| ||  | | ||| ||||| || || || ||||| |||||| |  ||
1184 ATCTCTGAGATGCTACGTCGTTTTGGCTGGATCGTCCCAGCTTACACTATGCCTGCCGAT 1243

1231 GCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCT 1290
     ||||| |||||||| ||||| ||||| || ||||| ||||||||||| |||||||||||
1244 GCACAGCACATCACGGTTCTGCGTGTTGTCATCAGGGAAGATTTCTCAAGAACACTCGCG 1303

1291 GAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGCTTCCTTCGAGA 1350
     |||||||||||    |||||   ||| ||| | | |||||| |||    | ||||| |
1304 GAGAGACTTGTTGCTGATATTTCGAAGGTGCTTCATGAGCTAGATACCTTGCCTTCCAAG 1363

1351 GTGATTCACAAAATATCACTTGGACAAGAGAAGAGTGAATCTAACAGCGATAACTTGATG 1410
      |                     |    ||||||  | |    |  |          |
1364 AT---------------------ATCTAAGAAGATGGGAATAGAAGGGATCGCGGAAAAT 1402

1411 GTCACGGTGAAGAAGAGCGATATCGACAAGCAGAGAGATATCATCACTGGCTGGAAGAAG 1470
     || | || ||||||||  || |  || |  | ||  ||  | ||   ||| |||| ||||
1403 GTAAAGGAGAAGAAGATGGAGAAGGAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAG 1462

1471 TTTGTCGCCGACAGGAAGAAGACGAGTGGTATCTGCTAA 1509
     |||||    || |||||||||| || |||| | ||||||
1463 TTTGTGAAGGAGAGGAAGAAGATGAATGGTGTGTGCTAA 1501


SULT 14
U49937
CUS        ATU49937              1681 bp    mRNA    linear   PLN 13-AUG-1998
FINITION   Arabidopsis thaliana glutamate decarboxylase (GAD2) mRNA, complete
           cds.
CESSION    U49937
RSION      U49937.1  GI:1236618
YWORDS     .
URCE       Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
FERENCE    1  (bases 1 to 1681)
AUTHORS    Zik,M., Arazi,T., Snedden,W.A. and Fromm,H.
TITLE      Two isoforms of glutamate decarboxylase in Arabidopsis are
           regulated by calcium/calmodulin and differ in organ distribution
JOURNAL    Plant Mol. Biol. 37 (6), 967-975 (1998)
MEDLINE    98363649
 PUBMED    9700069
FERENCE    2  (bases 1 to 1681)
AUTHORS    Zik,M., Arazi,T. and Fromm,H.
TITLE      Direct Submission
JOURNAL    Submitted (25-FEB-1996) Hillel Fromm, Plant Genetics, Weizmann
           Institute of Science, Rehovot 76100, Israel
ATURES              Location/Qualifiers
   source           1..1681
                    /organism="Arabidopsis thaliana"
                    /mol_type="mRNA"
                    /strain="Columbia"
                    /db_xref="taxon:3702"
                    /clone="5B2T7P"
                    /note="using EST found in GenBank Accession Number
                    T04804#"
   gene             1..1681
                    /gene="GAD2"
   CDS              1..1485
                    /gene="GAD2"
                    /function="calcium/calmodulin binding"
                    /note="glutamic acid decarboxylase; GAD; similar to A.
                    thaliana GAD encoded by GenBank Accession Numbers U10034
                    and U46665"
                    /codon_start=1
                    /product="glutamate decarboxylase"
                    /protein_id="AAC31617.1"
                    /db_xref="GI:1236619"
                    /translation="MVLTKTATNDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQI
                    IKDELMLDGNPRLNLASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNI
                    IARLFNAPLEESETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGA
```

```
                    NVQVCWEKFARYFEVELKEVNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFE
                    DVKRLNDLLVKKNEETGWNTPIHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHK
                    YGLVYAGIGWVVWRAAEDLPEELIFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRL
                    GFEGYKNVMENCIENMVVLKEGIEKTERFNIVSKDQGVPVVAFSLKDHSFHNEFEISE
                    MLRRFGWIVPAYTMPADAQHITVLRVVIREDFSRTLAERLVADISKVLHELDTLPSKI
                    SKKMGIEGIAENVKEKKMEKEILMEVIVGWRKFVKERKKMNGVC"
BASE COUNT      507 a    314 c    421 g    439 t
ORIGIN

  Query Match             57.7%;  Score 870.2;  DB 8;  Length 1681;
  Best Local Similarity   75.1%;  Pred. No. 5.7e-243;
  Matches 1110;  Conservative    0;  Mismatches  348;  Indels   21;  Gaps    1;

Qy          31 GACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCACTTCCTAGGTTC 90
               || |  || |||  | |||  |||| ||| || || || || ||| ||||||| | ||
Db          28 GATGAATCTGTCTGCACCATGTTCGGATCTCGCTATGTTCGCACTACACTTCCCAAGTAT 87

Qy          91 AAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTG 150
                ||||    || || ||||| || || || || || ||||||||||| || || ||||||
Db          88 GAGATTGGTGAGAATTCGATACCGAAAGACGCTGCATATCAGATCATAAAAGATGAGCTG 147

Qy         151 ATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCT 210
               |||||||| || || ||  || | ||| |||| || |||||||| ||||||||||| ||
Db         148 ATGCTTGATGGTAACCCGAGGCTTAACCTAGCTTCGTTTGTGACTACATGGATGGAACCA 207

Qy         211 GAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTAC 270
               |||||||| ||||||||||||  ||| |||||||||||||| ||||| ||||| ||||||
Db         208 GAGTGTGACAAACTCATCATGGACTCTATCAACAAGAACTACGTTGATATGGATGAGTAC 267

Qy         271 CCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATCTATTCAATGCA 330
               || ||||| || || || |||||||||||||| ||||| || || |  || ||||||||
Db         268 CCTGTCACAACTGAGCTCCAGAACCGATGTGTAAACATTATAGCTCGACTGTTCAATGCG 327

Qy         331 CCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATA 390
               ||  | || ||  | ||||| || || |||||||| || ||||| || || || |||||
Db         328 CCACTCGAGGAATCTGAGACGGCGGTGGGAGTAGGGACAGTTGGTTCTTCAGAAGCCATC 387

Qy         391 ATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAA 450
               ||||| ||||| |||||||||||  | |||||||||||||| ||||| ||||| || |||
Db         388 ATGTTAGCCGGATTGGCCTTCAAAAGAAAATGGCAGAACAAACGCAAGGCTGAGGGTAAA 447

Qy         451 CCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTC 510
               |||   || ||||||||||||||||| ||||||||||||||||| || ||||||||||||
Db         448 CCCTATGACAAACCCAACATTGTCACTGGAGCCAATGTTCAAGTTTGCTGGGAGAAATTC 507

Qy         511 GCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATG 570
               ||| ||||||| ||||| || || |||||||| ||  | |||||||| ||||| ||||||
Db         508 GCTCGGTACTTCGAGGTGGAGCTAAAGGAAGTAAACCTAAGTGAAGGTTACTACGTGATG 567

Qy         571 GACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTT 630
               || ||  |  |||| |  || ||||| || |||||||| || ||||| || | |||  |
Db         568 GATCCAGACAAAGCAGCAGAAATGGTAGACGAGAACACAATCTGTGTCGCAGCCATATTG 627

Qy         631 GGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAA 690
               || ||||| || || || || |||||||| || ||||   | || ||  |  | ||| |
Db         628 GGATCCACACTCAACGGTGAGTTCGAAGACGTGAAACGTCTCAATGACTTGCTAGTCAAG 687

Qy         691 AAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTC 750
               || ||| | || || || ||| | ||||| |||||||||||||| ||||||||||| |||
Db         688 AAAAACGAGGAGACTGGTTGGAACACACCGATCCACGTGGATGCAGCAAGTGGAGGGTTC 747

Qy         751 ATTGCACCGTTTTTGTATCCGGAATTGGAATGGACTTTAGACTTCCCTTGGTGAAGAGT 810
               || || |||||| | ||||| ||||| |||||||||||||||||||| ||||| ||||||
Db         748 ATAGCTCCGTTTATCTATCCTGAATTAGAATGGACTTTAGACTTCCTTTGGTTAAGAGT 807

Qy         811 ATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGG 870
               ||||| ||||||||||||||||||||||| || || || || |||||||||||  | |||
Db         808 ATCAACGTGAGTGGTCACAAGTATGGACTGGTCTATGCTGGTATTGGTTGGGTCGTGTGG 867

Qy         871 AGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGAC 930
               ||      ||||||||||||||| || || ||||| ||||| |||||||||||||||||
Db         868 AGGGCAGCAGAGGATTTGCCTGAAGAGCTTATCTTTCATATTAATTATCTTGGTGCTGAT 927
```

```
 931 CAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTAC 990
     |||||||| || ||||||||||||||||| || || || ||| | |||||||||||||||
 928 CAACCCACTTTCACTCTCAATTTCTCCAAGGGATCGAGCCAAATTATTGCTCAATACTAC 987

 991 CAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAAT 1050
     || || || ||  | |    |||||| |||| |||||||||||||||||||| ||||||
 988 CAGCTCATTCGTCTTGGATTCGAGGGGTACAAAAATGTGATGGAGAATTGCATAGAGAAC 1047

1051 ATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGAC 1110
     ||| | || || |  |||||  | ||||| |||||  | |||||||| ||||||||||||
1048 ATGGTGGTTCTCAAAGAAGGGATAGAGAAAACAGAGCGTTTCAACATAGTCTCAAAGGAC 1107

1111 GAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAA 1170
      | ||||||||| | || || |||||  | || ||    || |  ||||  ||||||||
1108 CAAGGAGTGCCAGTCGTAGCCTTCTCTCTCAAGGACCATAGTTTCCACAACGAGTTCGAG 1167

1171 ATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAAT 1230
     ||||| || ||||| ||  | | || |||||| || || || ||||| |||||| |  ||
1168 ATCTCTGAGATGCTACGTCGTTTTGGCTGGATCGTCCCAGCTTACACTATGCCTGCCGAT 1227

1231 GCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCT 1290
     ||||| |||||||| ||||| ||||| || ||||| ||||||||||| |||||||||||
1228 GCACAGCACATCACGGTTCTGCGTGTTGTCATCAGGGAAGATTTCTCAAGAACACTCGCG 1287

1291 GAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGCTTCCTTCGAGA 1350
     |||||||||||    |||||   ||| ||| | | |||||| |||    | ||||| |
1288 GAGAGACTTGTTGCTGATATTTCGAAGGTGCTTCATGAGCTAGATACCTTGCCTTCCAAG 1347

1351 GTGATTCACAAAATATCACTTGGACAAGAGAAGAGTGAATCTAACAGCGATAACTTGATG 1410
      |                     |    ||||||  | |    |  |          |
1348 AT---------------------ATCTAAGAAGATGGGAATAGAAGGGATCGCGGAAAAT 1386

1411 GTCACGGTGAAGAAGAGCGATATCGACAAGCAGAGAGATATCATCACTGGCTGGAAGAAG 1470
     || | || ||||||||  || |  || |  | ||  ||  | ||   ||| |||| ||||
1387 GTAAAGGAGAAGAAGATGGAGAAGGAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAG 1446

1471 TTTGTCGCCGACAGGAAGAAGACGAGTGGTATCTGCTAA 1509
     |||||    || |||||||||| || |||| | ||||||
1447 TTTGTGAAGGAGAGGAAGAAGATGAATGGTGTGTGCTAA 1485


SULT 15
428372
CUS        AF428372                1695 bp    mRNA    linear   PLN 18-OCT-2001
FINITION   Arabidopsis thaliana At1g65960/F12P19_12 mRNA, complete cds.
CESSION    AF428372
RSION      AF428372.1  GI:16226933
YWORDS     FLI_CDNA.
URCE       Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
FERENCE    1  (bases 1 to 1695)
AUTHORS    Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
           Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
           Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
           Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
           Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
           Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
           Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
           Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE      Arabidopsis cDNA clones
JOURNAL    Unpublished
FERENCE    2  (bases 1 to 1695)
AUTHORS    Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
           Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
           Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
           Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
           Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
           Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
           Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
           Davis,R.W., Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-OCT-2001) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
            Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J. Bowser,L.,
            Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
            Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
            Tang,C.C., Toriumi,M., Yamada,K., Yamamura, Y., Yu,G., Yu,S.,
            Davis,R.W.,  Theologis,A., and Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1..1695
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="1"
                     /clone="RAFL09-67-N18(R14793)"
                     /note="ecotype: Columbia"
     CDS             33..1517
                     /note="glutamate decarboxylase 2"
                     /codon_start=1
                     /product="At1g65960/F12P19_12"
                     /protein_id="AAL16302.1"
                     /db_xref="GI:16226934"
                     /translation="MVLTKTATNDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQI
                     IKDELMLDGNPRLNLASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNI
                     IARLFNAPLEESETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGA
                     NVQVCWEKFARYFEVELKEVNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFE
                     DVKRLNDLLVKKNEETGWNTPIHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHK
                     YGLVYAGIGWVVWRAAEDLPEELIFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRL
                     GFEGYKNVMENCIENMVVLKEGIEKTERFNIVSKDQGVPVVAFSLKDHSFHNEFEISE
                     MLRRFGWIVPAYTMPADAQHITVLRVVIREDFSRTLAERLVADISKVLHELDTLPSKI
                     SKKMGIEGIAENVKEKKMEKEILMEVIVGWRKFVKERKKMNGVC"
BASE COUNT      498 a    323 c    424 g    450 t
ORIGIN

  Query Match             57.7%;  Score 870.2;  DB 8;  Length 1695;
  Best Local Similarity   75.1%;  Pred. No. 5.7e-243;
  Matches 1110;  Conservative    0;  Mismatches  348;  Indels   21;  Gaps    1;

Qy         31 GACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCACTTCCTAGGTTC 90
              || |  || |||  | |||  |||| ||| || || || || ||| ||||||| | ||
Db         60 GATGAATCTGTCTGCACCATGTTCGGATCTCGCTATGTTCGCACTACACTTCCCAAGTAT 119

Qy         91 AAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTG 150
               ||||    || || ||||| || || || || |||||||||||||| || || ||||||
Db        120 GAGATTGGTGAGAATTCGATACCGAAAGACGCTGCGTATCAGATCATAAAAGATGAGCTG 179

Qy        151 ATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCT 210
              |||||||| || || ||| || | ||| |||| || |||||||| ||||||||||| ||
Db        180 ATGCTTGATGGTAACCCAAGGCTTAACCTAGCTTCTTTTGTGACTACATGGATGGAACCA 239

Qy        211 GAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTAC 270
              |||||||| ||||||||||||  ||| ||||| |||||||| ||||| ||||| ||||||
Db        240 GAGTGTGACAAACTCATCATGGACTCTATCAATAAGAACTACGTTGATATGGATGAGTAC 299

Qy        271 CCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATCTATTCAATGCA 330
              || ||||| || || || |||||||||||||| ||||| || || |  || ||||||||
Db        300 CCTGTCACAACTGAGCTCCAGAACCGATGTGTAAACATTATAGCTCGACTGTTCAATGCG 359
```

```
         331 CCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATA 390
             ||  | || ||  | ||||| || || ||||||||| || ||||| || || || |||||
         360 CCACTCGAGGAATCTGAGACGGCGGTGGGAGTAGGGACAGTTGGTTCTTCAGAAGCCATC 419

         391 ATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAA 450
             ||||| ||||| |||||||||||  | |||||||||||||| ||||| ||||| || |||
         420 ATGTTAGCCGGATTGGCCTTCAAAAGAAAATGGCAGAACAAACGCAAGGCTGAGGGTAAA 479

         451 CCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTC 510
             |||   || |||||||||||||||||||||||||||||||||||| || ||||||||||||
         480 CCCTATGACAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTTTGCTGGGAGAAATTC 539

         511 GCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATG 570
             ||| ||||||| ||||| || || |||||||| ||  | ||||||| ||||| ||||||
         540 GCTCGGTACTTCGAGGTGGAGCTAAAGGAAGTAAACCTAAGTGAAGGTTACTACGTGATG 599

         571 GACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTT 630
             || ||  |  |||| |  || ||||| || |||||||| || ||||| || | |||  |
         600 GATCCAGACAAAGCAGCAGAAATGGTAGACGAGAACACAATCTGTGTCGCAGCCATATTG 659

         631 GGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAA 690
             || ||||| || || || || |||||||| || ||||   | || ||  |  | ||| |
         660 GGATCCACACTCAACGGTGAGTTCGAAGACGTGAAACGTCTCAATGACTTGCTAGTCAAG 719

         691 AAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTC 750
             || ||| | || || || || | ||||| |||||||||||||| ||||||||||| |||
         720 AAAAACGAGGAGACTGGTTGGAACACACCGATCCACGTGGATGCAGCAAGTGGAGGGTTC 779

         751 ATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGT 810
             || || |||||| | ||||| ||||| |||||||||||||| ||||| ||||| |||||
         780 ATAGCTCCGTTTATCTATCCTGAATTAGAATGGGACTTTAGGCTTCCTTTGGTTAAGAGC 839

         811 ATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGG 870
             ||||| ||||||||||||||||||||||| || || || || |||||||||||  | |||
         840 ATCAACGTGAGTGGTCACAAGTATGGACTAGTCTATGCTGGTATTGGTTGGGTCGTGTGG 899

         871 AGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGAC 930
             ||      ||||||||| ||||| || || ||||| ||||| |||||||||||||||||
         900 AGGGCAGCAGAGGATTTACCTGAAGAGCTTATCTTTCATATTAATTATCTTGGTGCTGAT 959

         931 CAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTAC 990
             |||||||| || ||||||||||||||||| || || || ||| | |||||||||||||||
         960 CAACCCACTTTCACTCTCAATTTCTCCAAGGGATCGAGCCAAATTATTGCTCAATACTAC 1019

         991 CAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAAT 1050
             || || || ||  | ||   |||||| |||| |||||||||||||||||||| ||||||
        1020 CAGCTCATTCGTCTTGGATTCGAGGGGTACAAAAATGTGATGGAGAATTGCATAGAGAAC 1079

        1051 ATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGAC 1110
             ||| | || || |  |||||  | ||||| |||||  | ||||||| ||||||||||||
        1080 ATGGTGGTTCTCAAAGAAGGTATAGAGAAAACAGAGCGTTTCAACATAGTCTCAAAGGAC 1139

        1111 GAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAA 1170
              | ||||||||| | ||||| |||||  | || ||    || |  ||||   ||||||||
        1140 CAAGGAGTGCCAGTCGTCGCCTTCTCTCTCAAGGACCATAGTTTCCACAACGAGTTCGAG 1199

        1171 ATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAAT 1230
             ||||| || ||||| ||  | | ||| ||||| || || || ||||| ||||||| |  ||
        1200 ATCTCTGAGATGCTACGTCGTTTTGGCTGGATCGTCCCAGCTTACACTATGCCTGCCGAT 1259

        1231 GCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCT 1290
             ||||| |||||||| |||||  ||||| || ||||| ||||||||||| |||||||||||
        1260 GCACAGCACATCACGGTTCTGCGTGTTGTCATCAGGGAAGATTTCTCAAGAACACTCGCG 1319

        1291 GAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGCTTCCTTCGAGA 1350
             |||||||||||    |||||   ||| ||| | | |||||| |||    | ||||| |
        1320 GAGAGACTTGTTGCTGATATTTCGAAGGTGCTTCATGAGCTAGATACCTTGCCTTCCAAG 1379

        1351 GTGATTCACAAAATATCACTTGGACAAGAGAAGAGTGAATCTAACAGCGATAACTTGATG 1410
              |                     |    ||||||  | |    |  |          |
        1380 AT---------------------ATCTAAGAAGATGGGAATAGAAGGGATCGCGGAAAAT 1418

        1411 GTCACGGTGAAGAAGAGCGATATCGACAAGCAGAGAGATATCATCACTGGCTGGAAGAAG 1470
             || | || ||||||||  || |  || |  | ||  ||  | ||   ||| |||| ||||
Db      1419 GTAAAGGAGAAGAAGATGGAGAAGGAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAG 1478

Qy      1471 TTTGTCGCCGACAGGAAGAAGACGAGTGGTATCTGCTAA 1509
             |||||    || |||||||||| || |||| | ||||||
Db      1479 TTTGTGAAGGAGAGGAAGAAGATGAATGGTGTGTGCTAA 1517

Search completed: October 22, 2003, 12:44:29
Job time : 5522 secs
```

```
                     GenCore version 5.1.6
              Copyright (c) 1993 - 2003  Compugen Ltd.

M nucleic - nucleic search, using sw model

un on:         October 22, 2003, 10:16:42 ; Search time 2998 Seconds
                                            (without alignments)
                                            12233.307 Million cell updates/sec

itle:          US-10-006-852-1
erfect score:  1509
equence:       1 atggtgctctcccacgccgt..........agacgagtggtatctgctaa 1509

coring table:  IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

earched:       22781392 seqs, 12152238056 residues

otal number of hits satisfying chosen parameters:      45562784

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :       EST:*
               1:  em_estba:*
               2:  em_esthum:*
               3:  em_estin:*
               4:  em_estmu:*
               5:  em_estov:*
               6:  em_estpl:*
               7:  em_estro:*
               8:  em_htc:*
               9:  gb_est1:*
               10:  gb_est2:*
               11:  gb_htc:*
               12:  gb_est3:*
               13:  gb_est4:*
               14:  gb_est5:*
               15:  em_estfun:*
               16:  em_estom:*
               17:  em_gss_hum:*
               18:  em_gss_inv:*
               19:  em_gss_pln:*
               20:  em_gss_vrt:*
               21:  em_gss_fun:*
               22:  em_gss_mam:*
               23:  em_gss_mus:*
               24:  em_gss_pro:*
               25:  em_gss_rod:*
               26:  em_gss_phg:*
               27:  em_gss_vrl:*
               28:  gb_gss1:*
               29:  gb_gss2:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                 SUMMARIES

               %
sult         Query
 No.  Score  Match Length DB  ID                       Description
--------------------------------------------------------------------------
   1  603.2   40.0   1168  11  AY103733                AY103733 Zea mays
   2  568.4   37.7    570   9  AV541881               AV541881 AV541881
   3    548   36.3    548   9  AV554189               AV554189 AV554189
   4  526.2   34.9    811  14  CB655721                CB655721 OSJNEc09G
   5  525.2   34.8    837  10  BF275438                BF275438 GA__Eb002
   6  517.4   34.3    819  13  BQ165651                BQ165651 EST611520
   7  509.2   33.7    821  10  BE040860                BE040860 OF12G05 O
   8    506   33.5    554   9  AV552358               AV552358 AV552358
   9  500.4   33.2    775  13  BU026594                BU026594 QHG17F24.
  10  496.6   32.9    786  14  CA813481                CA813481 CA48LN10I
  11  486.8   32.3    762  13  BQ971999                BQ971999 QHB9G09.y
  12  484.8   32.1    823  10  BF263784                BF263784 HV_CEa000
  13  483.2   32.0    769  13  BU025918                BU025918 QHG12C07.
  14  482.8   32.0    750  13  BU023643                BU023643 QHF12C03.
  15    482   31.9    746  13  BU025264                BU025264 QHF8J20.y
  16    482   31.9    746  13  BU027680                BU027680 QHG7B13.y
  17  481.6   31.9    752  13  BU026039                BU026039 QHG12O09.
  18  480.4   31.8    746  13  BU025649                BU025649 QHG10J06.
  19  479.6   31.8    744  13  BU027771                BU027771 QHG7K14.y
  20  479.6   31.8    790  10  BG647367                BG647367 EST508986
  21    479   31.7    698  14  CD484071                CD484071 atr01-6ms
  22    479   31.7    744  13  BU025655                BU025655 QHG10K09.
  23    479   31.7    744  13  BU025839                BU025839 QHG11L03.
  24    479   31.7    744  13  BU025893                BU025893 QHG11P15.
  25  476.4   31.6    744  13  BU026030                BU026030 QHG12N09.
  26  475.2   31.5    721   9  AW622545               AW622545 EST313345
  27  475.2   31.5    746  13  BU026554                BU026554 QHG17D12.
  28  473.8   31.4    786  13  BU027739                BU027739 QHG7H16.y
  29  472.2   31.3    746  13  BU027486                BU027486 QHG5P23.y
  30  471.4   31.2    805  14  CB622579                CB622579 OSIIEa09H
  31  470.4   31.2    805  14  CB618146                CB618146 OSIIEa01N
  32  470.2   31.2    745  13  BU025328                BU025328 QHF8O15.y
  33    470   31.1    744  13  BU025929                BU025929 QHG12D11.
  34    468   31.0    732  13  BU026249                BU026249 QHG14D05.
  35  467.4   31.0    742  13  BU025333                BU025333 QHF8P03.y
  36  467.4   31.0    781  13  BU025900                BU025900 QHG12A05.
  37  466.6   30.9    721  13  BU027749                BU027749 QHG7I10.y
  38  465.8   30.9    726  13  BU024451                BU024451 QHF3g03.y
  39  465.4   30.8    721  10  BG589644                BG589644 EST497486
  40  465.2   30.8    726  13  BU027466                BU027466 QHG5N09.y
  41  463.8   30.7    743  13  BU025776                BU025776 QHG11F24.
  42  463.2   30.7    742  13  BQ916451                BQ916451 QHB17O10.
  43  462.6   30.7    727  13  BU023920                BU023920 QHF14N08.
  44  462.2   30.6    715  13  BU025894                BU025894 QHG11P17.
  45  461.8   30.6    737  13  BU026335                BU026335 QHG14O10.

                              ALIGNMENTS

RESULT 1
AY103733
LOCUS       AY103733               1168 bp    mRNA    linear   HTC 16-OCT-2002
DEFINITION  Zea mays PCO153431 mRNA sequence.
ACCESSION   AY103733
VERSION     AY103733.1  GI:21206811
KEYWORDS    HTC.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1168)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 1168)
  AUTHORS   Coe,E.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
```

```
        maize cDNA sequences is either Virginia Walbot, Stanford or Pat
        Schnable, Iowa State, then clones may be requested from ZmDB:
        www.zmdb.iastate.edu.
EATURES              Location/Qualifiers
  source             1..1168
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:638900"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
ASE COUNT      275 a    332 c    345 g    216 t
RIGIN

Query Match             40.0%;  Score 603.2;  DB 11;  Length 1168;
Best Local Similarity   74.2%;  Pred. No. 1.9e-165;
Matches  775;  Conservative    0;  Mismatches  268;  Indels    1;  Gaps    1;

          6 GCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCACGTTA 65
            || | |   | |||    |||    |||| | | ||   |||||| ||||| || || ||
        118 GCACGCGAGCTCCGGCCGGGACGACGACGACGCGGTGGCCTCCACCTTCGCGTCGCGCTA 177

         66 CGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGC 125
            ||| ||       || ||   ||| |  |||||||||    ||||| ||   ||| |||||
        178 CGTGCGCGAGCGCCTGCCGCGGTACCGGATGCCGGAGCGGTCGATCCCGCGGGAGGCGGC 237

        126 GTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTC 185
            | | |||||||||| ||||||||| ||||| ||||| || || ||  ||||| | || ||
        238 GCAGCAGATCATCAGCGACGAGCTCATGCTGGACGGCAACCCGCGCCTGAACCTGGCGTC 297

        186 CTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAA 245
            ||| || || || ||||||||||| |||||  || || |||||||||  ||||||||||||
        298 CTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATCATGGACTCCATCAACAA 357

        246 GAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAA 305
            |||||| || |||||||||||||||||||||||||| || || ||||||| | || || ||
        358 GAACTACGTCGACATGGACGAGTACCCCGTCACCACTGAGCTCCAGAACCGCTGCGTAAA 417

        306 CATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGG 365
            |||||| || || || |||| || ||| |  | ||||  ||||| ||  | || || ||
        418 CATGATAGCTCACCTGTTCAACGCGCCGATCAAGGAGGACGAGACAGCTATTGGGGTCGG 477

        366 AACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCA 425
            ||| || ||||| ||||| || || ||| | || ||| |||| |||||| | ||||||||
        478 AACAGTGGGATCCTCGGAAGCGATCATGCTTGCAGGTCTGGCGTTCAAGAGGAAATGGCA 537

        426 GAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAA 485
            ||||||| | || |   | || |||||     || | ||||||||| |||||||| |||||
        538 GAACAAGAGGAAGGAGCAGGGGAAACCGTGTGACAGACCCAACATCGTCACCGGTGCCAA 597

        486 TGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAA 545
             ||||| || || |||||||| || || ||||| |||||||| || || ||||| || ||
        598 CGTTCAGGTCTGCTGGGAGAAGTTTGCAAGGTATTTTGAGGTGGAGCTGAAGGAGGTTAA 657

        546 ATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAA 605
              |    ||||| ||||| || ||||||||     | |||||||| |||| || |||||
        658 GCTATCAGAAGGGTACTACGTCATGGACCCCGTCAAGGCTGTTGAGATGGTGGACGAGAA 717

        606 CACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAA 665
            ||| || ||||| ||||  ||  | || ||  ||||| |  ||||| |||||| ||||| ||
        718 CACTATCTGTGTCGCGGCGATCTTGGGGTCGACTCTCACCGGAGAGTTCGAGGATGTCAA 777

        666 ACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCA 725
             |  |||||||| || ||  |    || |||||||| ||||| || |||||    || |||||
        778 GCAGTTGAACGACCTTCTTACAGAGAAGAACAAGGAAACTGGGTGGGACGTGCCTATCCA 837

        726 CGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGA 785
```

```
                ||| || || || || ||||| ||||| || || ||  | || || ||   | || |||||
Db          838 CGTCGACGCCGCGAGCGGAGGGTTCATAGCGCCTTTCCTCTACCCTGAGCTCGAGTGGGA 897

Qy          786 CTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTA 845
                ||| || || ||  |||||||||| ||||||||| || || |||||||| || || || ||
Db          898 CTTCAGGCTGCCTCTGGTGAAGAGCATCAATGTCAGCGGGCACAAGTACGGCCTCGTCTA 957

Qy          846 CGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTT 905
                | | ||  |||| ||||| |||||| |  |||| || |||| ||||||||| ||||||||
Db          958 CCCTGGTGTTGGCTGGGTCATCTGGCG-GACAAGGATGATCTGCCTGAGGAGCTCATCTT 1016

Qy          906 CCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTC 965
                ||||||||| || || ||  | ||||| || || || || ||||| ||||||||||||||
Db         1017 CCATATCAACTACCTGGGGACGGACCAGCCTACGTTCACGCTCAACTTCTCCAAAGGTTC 1076

Qy          966 AAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAA 1025
                  ||||  ||||||| || ||||| ||||| || ||  | |||  |||||| ||||||||
Db         1077 CTGTCAGATCATTGCGCAGTACTATCAACTCATTCGCCTAGGCTTCGAGGGGTACAGAAA 1136

Qy         1026 TGTGATGGAGAATTGCAGAGAGAA 1049
                  | |||||||| |||   |||||
Db         1137 CATCATGGAGAACTGCCAGGAGAA 1160


RESULT 2
AV541881
LOCUS       AV541881                 570 bp    mRNA    linear   EST 07-SEP-2000
DEFINITION  AV541881 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ172g07F 3', mRNA sequence.
ACCESSION   AV541881
VERSION     AV541881.1  GI:8703640
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
   PUBMED   10907847
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1..570
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ172g07F"
                     /tissue_type="roots"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      161 a    104 c    156 g    149 t
ORIGIN

  Query Match             37.7%;  Score 568.4;  DB 9;  Length 570;
  Best Local Similarity   99.8%;  Pred. No. 2e-155;
  Matches  569;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          351 CGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTT 410
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 CGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTT 60

Qy          411 CAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACAT 470
```

```
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     61 CAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACAT 120

    471 TGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGA 530
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    121 TGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGA 180

    531 ACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGA 590
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    181 ACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGA 240

    591 TATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGA 650
        |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
    241 TATGGTTGATGAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGA 300

    651 ATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATG 710
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    301 ATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATG 360

    711 GGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCC 770
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    361 GGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCC 420

    771 GGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAA 830
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    421 GGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAA 480

    831 GTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCC 890
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    481 GTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCC 540

    891 TGAGGAACTCATCTTCCATATCAATTATCT 920
        ||||||||||||||||||||||||||||||
    541 TGAGGAACTCATCTTCCATATCAATTATCT 570


SULT 3
554189
CUS          AV554189                 548 bp    mRNA    linear   EST 06-SEP-2000
FINITION     AV554189 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
             cDNA clone RZ83g06R 5', mRNA sequence.
CESSION      AV554189
RSION        AV554189.1  GI:8725602
YWORDS       EST.
URCE         Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
             ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
FERENCE      1  (bases 1 to 548)
AUTHORS      Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE        A large scale analysis of cDNA in Arabidopsis thaliana: Generation
             of 12,028 non-redundant expressed sequence tags from normalized and
             size-selected cDNA libraries
JOURNAL      DNA Res. 7, 175-180 (2000)
MEDLINE      20363093
 PUBMED      10907847
MMENT        Contact: Erika Asamizu
             The First Laboratory for Plant Gene Research
             Kazusa DNA Research Institute
             Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
             Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
ATURES                Location/Qualifiers
   source             1. .548
                      /organism="Arabidopsis thaliana"
                      /mol_type="mRNA"
                      /strain="Columbia"
                      /db_xref="taxon:3702"
                      /clone="RZ83g06R"
                      /tissue_type="roots"
                      /clone_lib="Arabidopsis thaliana roots Columbia"
                      /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                      XhoI"
```

```
BASE COUNT      164 a    114 c    132 g    138 t
ORIGIN

  Query Match             36.3%;  Score 548;  DB 9;  Length 548;
  Best Local Similarity   100.0%;  Pred. No. 2e-149;
  Matches  548;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        849 AGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCA 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCA 60

Qy        909 TATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAG 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAG 120

Qy        969 TCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGT 1028
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGT 180

Qy       1029 GATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAG 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAG 240

Qy       1089 GTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAG 1148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAG 300

Qy       1149 CAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCC 1208
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCC 360

Qy       1209 GGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGA 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGA 420

Qy       1269 AGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGA 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGA 480

Qy       1329 GCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAGAAGAGTGA 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAGAAGAGTGA 540

Qy       1389 ATCTAACA 1396
              ||||||||
Db        541 ATCTAACA 548


RESULT 4
CB655721
LOCUS       CB655721                811 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  OSJNEc09G09.f OSJNEc Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEc09G09 5', mRNA sequence.
ACCESSION   CB655721
VERSION     CB655721.1  GI:29659446
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 811)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
```

```
          Fax: 520 621 9288
          Email: http://genome.arizona.edu
          PCR PRimers
          FORWARD: gta aaa cga cgg cca gtg
          BACKWARD: gga aac agc tat gac cat g
          Plate: 09  row: G  column: 09
          Seq primer: gta aaa cga cgg cca gtg.
ATURES               Location/Qualifiers
   source            1..811
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEc09G09"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEc"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 6 hrs after innoculation with Rice Blast (C9240-1)"
SE COUNT      182 a    236 c    244 g    149 t
IGIN

Query Match             34.9%;  Score 526.2;  DB 14;  Length 811;
Best Local Similarity   78.1%;  Pred. No. 6.3e-143;
Matches  633;  Conservative    0;  Mismatches  178;  Indels    0;  Gaps    0;

      179 TAGCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCA 238
          | || || || || || ||||||||||| ||||| || || ||||||   | |||||
        1 TCGCGTCGTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATCCAGGCCTCCG 60

      239 TCAACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGAT 298
          ||||||||||||| || ||||||||||||||||||||||||||||||||| ||||||||||
       61 TCAACAAGAACTACGTCGACATGGACGAGTACCCCGTCACCACCGAACTCCAGAACCGAT 120

      299 GTGTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCG 358
          ||||||||||||||||||| || ||||||||| || ||  |||  ||  | || |||||||
      121 GTGTGAACATGATTGCACACCTCTTCAATGCTCCTCTAGGGGACTCTGAAACGGCCGTCG 180

      359 GAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTA 418
          |||| || || || || || || ||||||||| ||| | ||||||||||||||||||| | |
      181 GAGTCGGCACTGTCGGCTCGTCTGAGGCCATCATGCTCGCCGGTTTGGCCTTCAAGAGGA 240

      419 AATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCG 478
            ||||||||||||   || || |  |||||| ||   ||| || || |||||||||||||
      241 GGTGGCAGAACAAGATGAAGGCAGCCGGCAAGCCATGCGACAAGCCTAACATTGTCACCG 300

      479 GAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGG 538
          | |||||||| ||||| || |||||||| |||||  | ||||| |||||||||| || ||||
      301 GCGCCAATGTCCAAGTTTGCTGGGAGAAGTTCGCGCGATACTTCGAGGTTGAGCTCAAGG 360

      539 AAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTG 598
          |||||||  |||||||| || ||||| || ||||||||||     | || || |||||||| |
      361 AAGTGAAGCTGAGTGACGGCTACTACGTCATGGACCCAGCTAAGGCCGTGGATATGGTCG 420

      599 ATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAG 658
          | ||||||||||| || || ||||  || || || || || || || || || ||||| |
      421 ACGAGAACACCATCTGCGTCGCGGCGATCCTCGGGTCGACGCTGAACGGGGAGTTCGAGG 480

      659 ATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACAC 718
          | || || ||   | ||||||||  |   | | ||||||   ||||| || ||||| || |
      481 ACGTGAAGCTGCTCAACGATCTGCTCACCAAGAAGAACGCTGAAACAGGCTGGGACACGC 540

      719 CAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGG 778
          | ||||||||||| ||||| || || || |||| || |||  |||| |||||  |||
      541 CGATCCACGTGGACGCGGCGAGCGGCGGGTTCATCGCGCCGTTCCTGTACCCGGAGCTGG 600

      779 AATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGAC 838
          | ||||||||   | || ||  ||||||||| | ||||| ||||| || |||||||| || |
      601 AGTGGGACTTCCGGCTGCCGCTGGTGAAGAGCATCAACGTGAGCGGGCACAAGTACGGCC 660

      839 TTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAAC 898
          | || ||||| ||||| || |||   |||||||| | ||| |||||| |||||||||| |
```

```
Db        661 TCGTCTACGCCGGGATCGGGTGGTGCATCTGGAGGAGCAAGGAGGATCTGCCTGAGGAGC 720

Qy        899 TCATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCA 958
              |||||||||| ||||| || || || || || |||| ||||||||| || ||| || |||||||
Db        721 TCATCTTCCACATCAACTACCTCGGCGCCGACCAGCCCACCTTCACCCTCAACTTCTCCA 780

Qy        959 AAGGTTCAAGTCAAGTCATTGCTCAATACTA 989
              | || |  || || ||||||||| || || ||
Db        781 AGGGGTTCAGCCAGGTCATTGCACAGTATTA 811


RESULT 5
BF275438
LOCUS       BF275438                 837 bp    mRNA    linear   EST 07-MAR-2001
DEFINITION  GA__Eb0024A19f Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA__Eb0024A19f, mRNA sequence.
ACCESSION   BF275438
VERSION     BF275438.1  GI:11206508
KEYWORDS    EST.
SOURCE      Gossypium arboreum
  ORGANISM  Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE   1  (bases 1 to 837)
  AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
            ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
  JOURNAL   Unpublished
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence start: 2
            High quality sequence stop: 718.
FEATURES             Location/Qualifiers
     source          1..837
                     /organism="Gossypium arboreum"
                     /mol_type="mRNA"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Eb0024A19f"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      244 a    171 c    198 g    223 t      1 others
ORIGIN

  Query Match             34.8%;  Score 525.2;  DB 10;  Length 837;
  Best Local Similarity   78.9%;  Pred. No. 1.3e-142;
  Matches  637;  Conservative    0;  Mismatches  169;  Indels    1;  Gaps    1;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              |||||| | ||  |  | | ||| ||  | || ||||| | |||||||| || || || 
Db         32 ATGGTGTTATCTAAGACAGCATCTGAAACTGATGTCTCAGTTCACTCCACTTTTGCCTCT 91

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              ||||| ||||| || ||||| || |||||||| ||||| ||||||| || || || || 
Db         92 CGTTATGTCCGAACCTCACTGCCCAGGTTCAAAATGCCAGAAAACTCCATACCAAAAGAG 151

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              || || |||||||||||||| || || ||||||||||||| | || ||| ||||||||| | 
Db        152 GCTGCTTATCAGATCATCAATGATGAACTGATGCTTGATGGCAACCCAAGGTTGAACCTT 211

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
```

```
        ||||| ||||| || ||||||||||||||||| |||||||| || || |||  |||||||
    212 GCCTCTTTTGTTACTACATGGATGGAGCCTGAATGTGATAAGCTTATAATGGACTCCATC 271

    241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
        ||||||||||| ||||||||||| || ||||| |||||||| || |||||||| || |||
    272 AACAAGAACTACGTTGACATGGATGAATACCCTGTCACCACTGAGCTTCAGAATCGTTGT 331

    301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
        ||||||||||| |||||||| ||||| ||||| || | |||| | ||| | || || ||
    332 GTGAACATGATAGCACATCTTTTCAACGCACCATTGGGAGAGTCAGAGGCAGCAGTTGGT 391

    361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
        || ||||| ||||| ||||||||||| || |||||||| ||  | || |||||| | |||
    392 GTGGGAACAGTTGGGTCATCGGAGGCAATCATGTTGGCTGGCCTTGCTTTCAAGAGAAAA 451

    421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
        ||||| |||||||| || ||||||||||||||    || || || || || || || |||
    452 TGGCAAAACAAGCGTAAGGCTGAAGGCAAACCTTATGAAAAGCCGAATATCGTGACTGGA 511

    481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
        ||||||||||| || || ||||||||||| || ||||||||||| || ||  | || |||
    512 GCCAATGTTCAGGTTTGCTGGGAGAAATTTGCAAGGTACTTTGAAGTGGAGTTGAAAGAA 571

    541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
        |||||  | || ||||| ||||||||||||||||||    |||| || || |||||||||
    572 GTGAAGCTTAGAGAAGGGTACTATGTGATGGACCCTGTCAAAGCAGTGGAAATGGTTGAT 631

    601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
        || |||||||| ||||| || |  ||  | ||||| || || ||||||||||| ||||||
    632 GAAAACACCATCTGTGTAGCTGCAATCTTGGGTTCAACCCTCAATGGAGAATTTGAAGAT 691

    661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
        || || ||  | ||||||||| ||  | |||||| ||| |||||| ||||||||||| ||
    692 GTCAAGCTTCTAAACGATCTTTTAATTGAAAAG-ACANAGAAACTGGATGGGATACCCCC 750

    721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
        || || || ||||| || ||||| ||||| ||||| || |||||||| || ||| | ||
    751 ATTCATGTTGATGCAGCTAGTGGTGGATTTATTGCGCCATTTTTGTACCCAGAACTTGAG 810

    781 TGGGACTTTAGACTTCCCTTGGTGAAG 807
        |||||||| |    ||||   ||||||
    811 TGGGACTTCAAGGCTCCCCCTGTGAAG 837
```

```
SULT 6
165651
CUS        BQ165651                819 bp    mRNA    linear   EST 25-APR-2002
FINITION   EST611520 KVKC Medicago truncatula cDNA clone pKVKC-10G4, mRNA
           sequence.
CESSION    BQ165651
RSION      BQ165651.1  GI:20308272
YWORDS     EST.
URCE       Medicago truncatula (barrel medic)
ORGANISM   Medicago truncatula
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
           ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
           Medicago.
FERENCE    1  (bases 1 to 819)
AUTHORS    VandenBosch,K., Endre,G., Silverstein,K, Town,C.D., Van Aken,S.,
           Utterback,T., Cheung,F. and Fraser,C.M.
TITLE      The Medicago truncatula 'kiloclone' set; ESTs selected and
           re-arrayed from various libraries
JOURNAL    Unpublished
MMENT      Contact: VandenBosch K
           Department of Plant Biology
           University of Minnesota
           220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
           Tel: 612 624 2755
           Fax: 612 625 1738
           Email: kvandenb@cbs.umn.edu
           TIGR sequence name: MTNAJ76TK Alias Clone name:MHAM-57I2 More
           information is available at: www.medicago.org
```

```
           Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1..819
                     /organism="Medicago truncatula"
                     /mol_type="mRNA"
                     /cultivar="A17"
                     /db_xref="taxon:3880"
                     /clone="pKVKC-10G4"
                     /tissue_type="mixed tissues"
                     /dev_stage="various stages"
                     /lab_host="XLOLR"
                     /clone_lib="KVKC"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                     was directionally ligated into the Unizap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-Zap phage using Ex-assist
                     helper phage and propagated in XLOLR cells."
BASE COUNT      245 a    158 c    188 g    228 t
ORIGIN

  Query Match             34.3%;  Score 517.4;  DB 13;  Length 819;
  Best Local Similarity   79.4%;  Pred. No. 2.4e-140;
  Matches  625;  Conservative    0;  Mismatches  161;  Indels    1;  Gaps    1;

Qy          3 GGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCACG 62
              ||| |||||  |  | |  || ||||| || ||||| |||||||| || || || || ||
Db         31 GGTTCTCTCAAAGACAGCCTCCGAGTCTGATGTCTCTGTCCACTCAACCTTTGCTTCTCG 90

Qy         63 TTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGC 122
               || ||| | |||||||||||||| ||||||||||| ||  | || || || ||||| ||
Db         91 CTATGTCAGAACTTCACTTCCTAGATTCAAGATGCCAGAGGAGTCTATACCAAAGGATGC 150

Qy        123 GGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGC 182
               || || || || || ||||| ||  |||||||||| || || ||  | ||||| || ||
Db        151 AGCATACCAAATAATAAACGATGAATTGATGCTTGATGGAAACCCTAGATTGAATTTGGC 210

Qy        183 CTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAA 242
               ||||||||||| |||||||||||||||||||||||||||||||||||| | ||||| ||
Db        211 ATCCTTTGTGACAACATGGATGGAGCCTGAGTGTGATAAACTCATCATGGCTTCCATTAA 270

Qy        243 CAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGT 302
              ||||||||| ||||||||||| || ||||| |||||||| || || ||||| || |||||
Db        271 CAAGAACTACGTTGACATGGATGAATACCCAGTCACCACTGAGCTACAGAATCGGTGTGT 330

Qy        303 GAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGT 362
               |||||||| || ||||| |||||||||||  | ||||||   ||| | || || || ||
Db        331 TAACATGATAGCTCATCTTTTCAATGCACCACTTGAAGAGAATGAGGCTGCAGTTGGTGT 390

Qy        363 AGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATG 422
               || || ||||| ||||| ||||| |||||||| || || ||||| |||||  | || ||
Db        391 TGGCACTGTTGGCTCATCAGAGGCTATAATGTTAGCTGGATTGGCATTCAAAAGGAAGTG 450

Qy        423 GCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGC 482
              |||||||||  | |||   || || || ||    || ||||| ||||||||||| |||||
Db        451 GCAGAACAAAAGAAAACAAGAGGGAAAGCCTTATGACAAACCTAACATTGTCACTGGAGC 510

Qy        483 CAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGT 542
              |||||| || || |||||||||||||| || ||||||||||| || ||| | || || ||
Db        511 CAATGTCCAGGTTTGTTGGGAGAAATTTGCAAGGTACTTTGAAGTGGAATTGAAAGAGGT 570

Qy        543 GAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGA 602
              |||  | ||||||||||||||||| |||||||||  | ||||||| ||  |||| |||||
Db        571 GAAGCTAAGTGAAGGATACTATGTAATGGACCCTGCAAAAGCTGTGGAATTGGTAGATGA 630

Qy        603 GAACA-CCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATG 661
              ||||| ||||||||||||| |  || ||||||||||| || |||||||| ||||||||||
Db        631 GAACACCCATTTGTGTTGCTGCTATCCTTGGTTCCACACTAAATGGAGAGTTCGAAGATG 690

Qy        662 TTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAA 721
              |||||| ||| ||||||||| | || |||||||| || ||||| |||||||| || || |
Db        691 TTAAACGCTTAAACGATCTCCTTGTTGAAAAGAATAAGGAAACTGGATGGGACACTCCTA 750
```

```
       722 TCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAAT 781
           | || |||||||| || ||||| || |||||||| || ||| | ||||| ||| | || |
       751 TTCATGTGGATGCAGCTAGTGGTGGCTTCATTGCTCCATTTATTTATCCAGAACTTGAGT 810

       782 GGGACTT 788
           |||| ||
       811 GGGATTT 817


SULT 7
040860
CUS        BE040860                821 bp    mRNA    linear   EST 07-JUN-2000
FINITION   OF12G05 OF Oryza sativa cDNA 5' similar to glutamate decarboxylase
           1, mRNA sequence.
CESSION    BE040860
RSION      BE040860.1  GI:8336367
YWORDS     EST.
URCE       Oryza sativa
ORGANISM   Oryza sativa
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzeae; Oryza.
FERENCE    1  (bases 1 to 821)
AUTHORS    Bohnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferrea
           ,H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C.,
           Scara,G., Wheeler,M. and Zepeda,G.R.
TITLE      Functional Genomics of Plant Stress Tolerance
JOURNAL    Unpublished
MMENT      Contact: Michalowski,C.B.
           University of Arizona
           Bio Sciences West room 513, Tucson, AZ 85721, USA
           Tel: 520-621-7982
           Fax: 520-621-1697
           Email: cbm@u.arizona.edu
           An open reading frame exists.
           Insert Length: 1   Std Error: 0.00.
ATURES              Location/Qualifiers
   source           1..821
                    /organism="Oryza sativa"
                    /mol_type="mRNA"
                    /strain="Pokkali"
                    /db_xref="taxon:4530"
                    /tissue_type="entire plant"
                    /dev_stage="2 weeks"
                    /clone_lib="OF"
                    /note="1 week 150mM NaCl"
SE COUNT      185 a    231 c    254 g    149 t      2 others
IGIN

Query Match             33.7%;  Score 509.2;  DB 10;  Length 821;
Best Local Similarity   78.0%;  Pred. No. 6.3e-138;
Matches  613;  Conservative    0;  Mismatches  173;  Indels    0;  Gaps    0;

       132 GATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGT 191
           ||  |||||||||||||||||||| ||||| || || ||| ||||| | || || || ||
         5 GAGGATCAACGACGAGCTGATGCTGGACGGCAACCCGCGGCTGAACCTCGCGTCGTTCGT 64

       192 GACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTA 251
            || || ||||||||||| ||||| || || ||||||  | ||||| |||||||||||||
        65 CACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATCCAGGCCTCCGTCAACAAGAACTA 124

       252 TGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGAT 311
            || ||||||||||||||||||||||||||||| || |||||||||||||||||||||||
       125 CGTCGACATGGACGAGTACCCCGTCACCACCGAGCTCCAGAACCGATGTGTGAACATGAT 184

       312 TGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGT 371
           |||||| || |||||||| ||  |||  ||  | || || ||||||||||| || || ||
       185 TGCACACCTCTTCAATGCTCCTCTAGGGGACTCTGAAACGGCCGTCGGAGTCGGCACTGT 244

       372 TGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAA 431
            || || || |||||||| ||| | |||||||||||||||||| | |  |||||||||||
       245 CGGCTCGTCTGAGGCCATCATGCTCGCCGGTTTGGCCTTCAAGAGGAGGTGGCAGAACAA 304
```

```
Qy         432 GCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCA 491
               |   || || |  ||||| ||   ||| || || |||||||||||||| |||||||| ||
Db         305 GATGAAGGCAGCCGGCAAGCCATGCGACAAGCCTAACATTGTCACCGGCGCCAATGTCCA 364

Qy         492 AGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAG 551
               ||| || |||||||| |||||  | ||||| |||||||| || |||||||||||  ||||
Db         365 AGTTTGCTGGGAGAAGTTCGCGCGATACTTCGAGGTTGAGCTCAAGGAAGTGAAGCTGAG 424

Qy         552 TGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCAT 611
               ||| || ||||| || ||||||||     | || || |||||||| || |||||||||||
Db         425 TGACGGCTACTACGTCATGGACCCAGCTAAGGCCGTGGATATGGTCGACGAGAACACCAT 484

Qy         612 TTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTT 671
                || || ||||  || || || || || || || || ||||| || || || ||  |
Db         485 CTGCGTCGCGGCGATCCTCGGGTCGACGCTGAACGGGGAGTTCGAGGACGTGAAGCTGCT 544

Qy         672 GAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGA 731
                ||||||||  |   | | ||||||   ||||| || ||||| || || |||||||||||
Db         545 CAACGATCTGCTCACCAAGAAGAACGCTGAAACAGGCTGGGACACGCCGATCCACGTGGA 604

Qy         732 TGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAG 791
                ||||| || || || ||||| || |||||  |||| |||||  |||| ||||||||| |
Db         605 CGCGGCGAGCGGCGGGTTCATCGCGCCGTTCCTGTACCCGGAGCTGGAGTGGGACTTTCG 664

Qy         792 ACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGG 851
                || ||  |||||||||| ||||| ||||| || |||||||| || || || ||||| ||
Db         665 GCTGCCGCTGGTGAAGAGCATCAACGTGAGCGGGCACAAGTACGGGCTCGTCTACGCCGG 724

Qy         852 GATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATAT 911
               |||| | |||   ||||||   | ||| |||||| |||||||||| |||||||| || ||
Db         725 GATTCGGTGGTGCATCTGGGAGAGCAAGGAGGATCTGCCTGAGGAGCTCATCTTTCACAT 784

Qy         912 CAATTA 917
                | |||
Db         785 AACTTA 790


RESULT 8
AV552358
LOCUS       AV552358                554 bp    mRNA    linear   EST 06-SEP-2000
DEFINITION  AV552358 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
            cDNA clone RZ27f07R 5', mRNA sequence.
ACCESSION   AV552358
VERSION     AV552358.1  GI:8723771
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 554)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
   PUBMED   10907847
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1..554
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="RZ27f07R"
                     /tissue_type="roots"
                     /clone_lib="Arabidopsis thaliana roots Columbia"
```

```
                /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                XhoI"
SE COUNT      140 a    150 c    131 g    133 t
IGIN

Query Match             33.5%;  Score 506;  DB 9;  Length 554;
Best Local Similarity   100.0%;  Pred. No. 4.3e-137;
Matches  506;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

        1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       49 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 108

       61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      109 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 168

      121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      169 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 228

      181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      229 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 288

      241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      289 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 348

      301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      349 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 408

      361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      409 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 468

      421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      469 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 528

      481 GCCAATGTTCAAGTGTGTTGGGAGAA 506
          ||||||||||||||||||||||||||
      529 GCCAATGTTCAAGTGTGTTGGGAGAA 554


SULT 9
026594
CUS         BU026594                 775 bp    mRNA    linear   EST 23-AUG-2002
FINITION    QHG17F24.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
            clone QHG17F24, mRNA sequence.
CESSION     BU026594
RSION       BU026594.1  GI:22462114
YWORDS      EST.
URCE        Helianthus annuus (common sunflower)
ORGANISM    Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
FERENCE     1  (bases 1 to 775)
AUTHORS     Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
TITLE       Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
JOURNAL     Unpublished
MMENT       Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
```

```
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QHG17  row: F  column: 24.
FEATURES             Location/Qualifiers
     source          1..775
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHG17F24"
                     /lab_host="E.coli"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QH_EFGHJ sunflower RHA280
                     TAG_TISSUE=hulls
                     TAG_SEQ=GCTAGTCGGG"
BASE COUNT      224 a    140 c    191 g    219 t      1 others
ORIGIN

  Query Match             33.2%;  Score 500.4;  DB 13;  Length 775;
  Best Local Similarity   77.8%;  Pred. No. 2.3e-135;
  Matches  603;  Conservative    0;  Mismatches  172;  Indels    0;  Gaps    0;

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              || ||||| || || ||||| || ||||| |||||||||||  | |  |||||  | |||
Db          1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||| |||||    ||||||   ||||||||    || ||||||| |||||||| |||||
Db         61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              ||||| ||||| || |||||||||||||| ||| |||| ||||| || ||  | ||||||
Db        121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              ||||| |||||| ||||| ||||| |||||||| ||| |   ||||||| || |||||||||
Db        181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              || ||||| |||||||| || |  ||  | |||||||| || ||||| ||||| ||||||
Db        241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| |  || || || |||||| | || || ||      | |||||||||||||||||
Db        301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || || ||||| || |||||||||| || |||||||||| ||| | || ||||||| | |||
Db        361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||| ||  |  | || ||||| |||||||| ||||| ||||||||||||| ||||| |||
Db        421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || ||||| || ||||| ||||| || ||||| ||||||||| || |||||||| ||||||
Db        481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              ||||| | ||||||| ||||||||||||||| ||||| || || |||||||| |||||||||
Db        541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600
```

```
  961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
      ||||| |||||| | |||||||||||||| ||  | || || |||||   |||||| |||
  601 GGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

 1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
      |  | | | ||||| || ||   ||| |||   |  || |||| | ||||  | |||||
  661 AAGAGTATCATGGAAAACTGTNCAGAAAATGCAAATGTTCTAAAGAAAGGTTTGGAGAAA 720

 1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCT 1135
      || | | ||||||||||| |||| || ||  | || || |||||||| || ||||
  721 ACCGGACGGTTCAACATCATCTCTAAAGAACACGGCGTCCCACTTGTAGCCTTCT 775


SULT 10
813481
CUS        CA813481                786 bp    mRNA    linear   EST 11-APR-2003
FINITION   CA48LN10IF-E8 Cabernet Sauvignon Leaf - CA48LN Vitis vinifera cDNA
           clone CA48LN10IF-E8 5', mRNA sequence.
CESSION    CA813481
RSION      CA813481.1  GI:26262418
YWORDS     EST.
URCE       Vitis vinifera
ORGANISM   Vitis vinifera
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
           ; Vitaceae; Vitis.
FERENCE    1  (bases 1 to 786)
AUTHORS    Goes da Silva,F., Lim,H., Iandolino,A., Baek,J., Jones,K., Walker
           ,M.A. and Cook,D.R.
TITLE      Transcriptional responses of Vitis vinifera to infection by the
           bacterial pathogen Xylella fastidiosa
JOURNAL    Unpublished
MMENT      Contact: Doug Cook
           CAES Genome Facility
           UC Davis Department of Plant Pathology
           1 Shields Ave., DAvis, CA 95616, USA
           Tel: 530 754 6561
           Fax: 530 754 6617
           Email: drcook@ucdavis.edu
           Seq primer: GTTATCAGTCGACGGTACC.
ATURES              Location/Qualifiers
   source           1..786
                    /organism="Vitis vinifera"
                    /mol_type="mRNA"
                    /cultivar="Cabernet Sauvignon"
                    /db_xref="taxon:29760"
                    /clone="CA48LN10IF-E8"
                    /sex="hermaphrodite"
                    /dev_stage="Late season sample"
                    /lab_host="DH5alpha"
                    /clone_lib="Cabernet Sauvignon Leaf - CA48LN"
                    /note="Organ: Leaf; Vector: pDNR; Site_1: SfiI; Site_2:
                    SfiI; CA48LN is a cDNA library of Cabernet Sauvignon
                    leaves. The leaves were collected on September 20, 2001,
                    in Napa Valley, California, and represent leaves in late
                    season development. These leaves were asymptomatic and
                    verified to be non-infected with the bacterial pathogen,
                    Xylella fastidiosa, based on a diagnostic assay using PCR
                    and Xylella-specific primer pairs. cDNAs were made by
                    oligo-dT priming and directionally cloned. 5'and 3'
                    adaptors were used in cloning as follows:
                    5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                    5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Library was
                    constructed using the Clontech Creator SMART kit and
                    size-selected to contain the 0.5-3 kb size fraction."
SE COUNT      222 a    163 c    200 g    201 t
IGIN

Query Match             32.9%;  Score 496.6;  DB 14;  Length 786;
Best Local Similarity   77.8%;  Pred. No. 3.1e-134;
Matches  612;  Conservative    0;  Mismatches  174;  Indels    1;  Gaps    1;

     523 GAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAA 582
```

```
              ||||| ||| | |||||||||||  |||  || || ||||||||||||||||||    ||
Db          1 GAGGTGGAATTGAAGGAAGTGAAGCTGACGGATGGCTACTATGTGATGGACCCTGTCAAA 60

Qy        583 GCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTT 642
              || || || ||||| || ||||||||||| |||||||| |  || |||||||| || ||
Db         61 GCAGTGGAGATGGTGGACGAGAACACCATCTGTGTTGCTGCTATCCTTGGTTCTACCCTC 120

Qy        643 AATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAA 702
              ||||||||||||||||||||||||| |||||| ||||| |||    | || |||||||||
Db        121 AATGGAGAATTCGAAGATGTTAAACACTTGAATGATCTTTTGACAAAGAAAAACAAAGAA 180

Qy        703 ACCGGATGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTT 762
              || ||||||||||| |||||||| || ||||| |||||||| || || |||||||| |||
Db        181 ACAGGATGGGATACTCCAATCCATGTAGATGCAGCAAGTGGTGGCTTTATTGCACCCTTT 240

Qy        763 TTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGT 822
               |  | || ||  |||| ||||||||  |  | ||  | ||||| || || || || ||
Db        241 CTTCACCCAGAGCTGGAGTGGGACTTCCGCTTGCCACTTGTGAAAAGCATTAACGTTAGC 300

Qy        823 GGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAG 882
              || ||||||||| || |||||||| || || || ||||||||||| ||||| |  ||||||
Db        301 GGCCACAAGTACGGGCTTGTGTATGCTGGCATAGGTTGGGTGATTTGGAGGAGTAAAGAG 360

Qy        883 GATTTGCCTGAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTT 942
              || ||||| || || |||||||| |||||||| |||||||| || |||||||| |||||
Db        361 GACTTGCCCGATGAGCTCATCTTTCATATCAACTATCTTGGAGCCGACCAACCTACCTTC 420

Qy        943 ACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGA 1002
              || ||||| |||||||||||||| |||||||| |||||||| || || ||||| || ||
Db        421 ACCCTCAACTTCTCCAAAGGTTCTAGTCAAGTTATTGCTCAGTATTATCAACTCATTCGC 480

Qy       1003 TTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTA 1062
              |||||||| ||||| || || ||  | ||||| || ||   ||| |||   |  || ||
Db        481 TTGGGCCATGAGGGATATAGGAACATTATGGAAAACTGTCAAGAAAATGCAAGGGTGCTC 540

Qy       1063 AGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCA 1122
              |  |||||| | ||||||||||     |||||||| || || || ||  | |||||||||
Db        541 AAAGAAGGATTGGAGAAGACAGGGCACTTCAACATTGTGTCTAAAGAGAATGGAGTGCCA 600

Qy       1123 CTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATG 1182
               | || || ||||| |||||||| | ||||||||||| |||||||||  |  || |||||
Db        601 TTAGTAGCCTTCTCTTTGAAAGACAACAGCTGTCACAATGAGTTCGAGGTGTCAGACATA 660

Qy       1183 CTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATC 1242
              || ||  | ||||||||||| ||||| || || || || || ||  ||||||||||  |
Db        661 CTGCGACGCTATGGATGGATCGTGCCAGCTTATACCATACCGCCGGATGCACAACATGTG 720

Qy       1243 ACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTG 1302
              |||||||||||||| || || |||||||| |||||| | ||||| |||||| | || ||
Db        721 ACTGTTCTTCGTGTCGTCATAAGAGAAGA-TTCTCGCGCACACTTGCTGAGCGTCTGGTC 779

Qy       1303 ATCGATA 1309
              |  ||||
Db        780 ACAGATA 786


RESULT 11
BQ971999
LOCUS       BQ971999                762 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  QHB9G09.yg.ab1 QH_ABCDI sunflower RHA801 Helianthus annuus cDNA
            clone QHB9G09, mRNA sequence.
ACCESSION   BQ971999
VERSION     BQ971999.1  GI:22389520
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
  ORGANISM  Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
```

```
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
TITLE       Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
JOURNAL     Unpublished
)MMENT      Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QHB9  row: G  column: 09.
ATURES               Location/Qualifiers
  source             1..762
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA801"
                     /db_xref="taxon:4232"
                     /clone="QHB9G09"
                     /lab_host="E.coli"
                     /clone_lib="QH_ABCDI sunflower RHA801"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QH_ABCDI sunflower RHA801
                     TAG_TISSUE=hulls
                     TAG_SEQ=GCTAGTCGGG"
SE COUNT      224 a    138 c    190 g    210 t
IGIN

Query Match             32.3%;  Score 486.8;  DB 13;  Length 762;
Best Local Similarity   77.4%;  Pred. No. 2.3e-131;
Matches  590;  Conservative    0;  Mismatches  172;  Indels    0;  Gaps    0;

         361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
             || ||||| || || ||||| || ||||| |||||||||||  | || |||||  | |||
           1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

         421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
             ||||| |||||    ||||||   ||||||||    || ||||||| |||||||| ||||| 
          61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

         481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
             ||||| ||||| || |||||||||||||| ||| |||| ||||||| ||  | ||||||
         121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180

         541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
             ||||| ||||| ||||| ||||| |||||||| ||| |  ||||||| || |||||||||
         181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

         601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
             || ||||| |||||||| || |  ||  | |||||||| || ||||| ||||| ||||||
         241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

         661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
             ||||| |  || || || |||||| | || || ||     | ||||||||||||||||| 
         301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
             || || ||||| || ||||||||||| || |||||||| ||| | |  |||||| | |||
         361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
             ||||| ||  |  | || |||| |||||||| ||||| ||||||||||| ||||| |||
```

```
Db         421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

Qy         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
               || ||||| || ||||| ||||| || ||| ||||||||| || |||||||| ||||||
Db         481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

Qy         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
               ||||| || ||||| |||||||||||||| ||||| || || |||||||| |||||||||
Db         541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600

Qy         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
               ||||| ||||||   |||||||||||||| ||  | || || |||||   |||||| |||
Db         601 GGTTCTAGTCAAACAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

Qy        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
               |  |   | ||||| || ||   ||| |||   |  || |||| | ||||  | ||||| 
Db         661 AAGAGCATCATGGAAAACTGTCAAGAAAATGCAAATGTTCTAAAGAAAGGTTTGGAGAAA 720

Qy        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCA 1122
               || | | || |||||||| |||| || ||| | || || |||
Db         721 ACCGGACGGGTCAACATCATCTCTAAAGACAACGGCGTCCCA 762


RESULT 12
BF263784
LOCUS       BF263784                 823 bp    mRNA    linear   EST 23-OCT-2001
DEFINITION  HV_CEa0007K14f Hordeum vulgare seedling green leaf EST library
            HVcDNA0004 (Blumeria challenged) Hordeum vulgare subsp. vulgare
            cDNA clone HV_CEa0007K14f, mRNA sequence.
ACCESSION   BF263784
VERSION     BF263784.2  GI:13261070
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 823)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Wei,F., Begum,D.,
            Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi
            ,D.W., Fenton,R.D., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Blumeria infected incompatible (Mla13)
            seedling leaf cDNA library
  JOURNAL   Unpublished
COMMENT     On Nov 17, 2000 this sequence version replaced gi:11194778.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 571
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 721.
FEATURES             Location/Qualifiers
     source          1..823
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="CI16155 (Mla13)"
                     /db_xref="taxon:112509"
                     /clone="HV_CEa0007K14f"
                     /tissue_type="seedling green leaf"
                     /lab_host="TJC121"
                     /clone_lib="Hordeum vulgare seedling green leaf EST
                     library HVcDNA0004 (Blumeria challenged)"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1;
                     C.I. 16155 (Mla13) plants were greenhouse grown in the R
                     Wise lab at Iowa State University, Ames, IA; 7 day old
                     green seedlings were challenged with isolate A27 (AvrMla13
                     ) of Blumeria graminis f. sp. hordei, and leaves were
                     harvested 20 and 24 hr post-inoculation and snap frozen;
```

```
                    uninoculated leaves were harvested 20 hr post-inoculation
                    (Wei, Wise). In the TJ Close lab at the University of
                    California, Riverside, total RNA was prepared from each
                    sample pool, equal quantities of all three RNA pools were
                    combined, poly(A) RNA was purified from the mixture, one
                    cDNA library was made, and 1 million pfu were in vivo
                    excised to give pBluescript SK(-) cDNA phagemids (Choi,
                    Close). Phagemids were plated and picked at the Clemson
                    University Genomics Institute (CUGI) (Begum, Palmer,
                    Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
                    sequencing and sequence analysis were performed at CUGI
                    (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
                    The sequence has been trimmed to remove vector sequence
                    and contains a minimum of 100 bases of phred value 20 or
                    above. For more details on library preparation and
                    sequence analysis see
                    http://www.genome.clemson.edu/projects/barley. To order
                    this clone see http://www.genome.clemson.edu/orders Also
                    see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                    Genetically and physically anchored EST resources for
                    barley genomics. Barley Genetics Newsletter 31:29-30.
                    (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
ASE COUNT      209 a    187 c    249 g    175 t      3 others
IGIN

Query Match             32.1%;  Score 484.8;  DB 10;  Length 823;
Best Local Similarity   77.2%;  Pred. No. 9.2e-131;
Matches  602;  Conservative    0;  Mismatches  177;  Indels    1;  Gaps    1;

        100 GAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGAC 159
            ||||||||||| || ||||| |||||||| |||||||||| ||||||||||||||| |||
         43 GAAAACTCGATCCCCAAGGAGGCGGCGTACCAGATCATCAGCGACGAGCTGATGCTGGAC 102

        160 GGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGAT 219
            || || || ||| ||||| | || ||||| || || || ||||||||||| ||||| |
        103 GGCAACCCGCGGCTGAACCTGGCGTCCTTCGTCACCACCTGGATGGAGCCCGAGTGCGGC 162

        220 AAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTACCCCGTCACC 279
            || |||||||||  |||| ||||||||||||| || ||||||||||||||||||||||||
        163 AAGCTCATCATGGACTCCGTCAACAAGAACTACGTCGACATGGACGAGTACCCCGTCACC 222

        280 ACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATCTATTCAATGCACCGTTAGAA 339
            ||||| || |||||||| || || |||||||| || ||  | |||||||||||| | |
        223 ACCGAGCTCCAGAACCGTTGCGTAAACATGATAGCTCACTTGTTCAATGCACCGATCGGC 282

        340 GAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCC 399
            ||||  ||||| ||  |||||||    || || || || ||||| || || ||| | ||
        283 GAGGACGAGACAGCTATCGGAGTCTCGACGGTGGGGTCTTCGGAAGCAATCATGCTTGCA 342

        400 GGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGAT 459
            ||  |||| |||||| | || |||  ||||||    || |   | || || ||   |||
        343 GGCCTGGCGTTCAAGAGGAAGTGGGCGAACAAAATGAAGGAGCAGGGGAAGCCATGCGAC 402

        460 AAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTAC 519
            ||||| |||||||| || || || |||||||| || || ||||||||||| ||||||||
        403 AAACCTAACATTGTTACTGGTGCAAATGTTCAGGTTTGCTGGGAGAAATTTGCTAGGTAT 462

        520 TTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAA 579
            ||||| || ||| | ||||| || || |||| |||||| |||||||| ||||| ||| |
        463 TTTGAAGTGGAATTGAAGGAGGTCAAGTTGACTGAAGGGTACTATGTCATGGATCCTAAG 522

        580 CAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACT 639
             | |||||||| ||||| ||||||||||| || ||||| || | ||| || || || |||
        523 AAGGCTGTTGAAATGGTGGATGAGAACACTATATGTGTCGCCGCCATCCTGGGATCTACT 582

        640 CTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAA 699
            || | |||||| | ||||||||| ||||| ||||| || ||  | || || |||||||||
        583 CTCACTGGAGAGTACGAAGATGTCAAACTGTTGAATGACCTTCTTGTGGAGAAGAACAAA 642

        700 GAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCG 759
            ||||| || ||| |    || ||||| || ||||| || || |||||||| || || |||
        643 GAAACAGGGTGGAACGTGCCGATCCATGTTGATGCTGCCAGCGGAGGATTTATCGCTCCG 702
```

```
Qy        760 TTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTG 819
              ||| | || || ||  | ||||||||||| |  || || ||||||||||| ||||| |
Db        703 TTTCTTTAGCCTGAGCTTGAATGGGACTTCAAGCTACCATTGGTGAAGAGCATCAACGGT 762

Qy        820 AGTGGTCACAAGTA-TGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAA 878
              || || |||||||| ||  ||||||||| | ||| |||| ||||| || ||| | | |||
Db        763 AGGGGGCACAAGTATTGCCCTTGTGTACCCTGGGGTTGGATGGGTCATTTGGCGGAGCAA 822


RESULT 13
BU025918
LOCUS       BU025918                 769 bp    mRNA    linear   EST 23-AUG-2002
DEFINITION  QHG12C07.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
            clone QHG12C07, mRNA sequence.
ACCESSION   BU025918
VERSION     BU025918.1  GI:22461438
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
  ORGANISM  Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE   1  (bases 1 to 769)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QHG12  row: C  column: 07.
FEATURES             Location/Qualifiers
     source          1..769
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHG12C07"
                     /lab_host="E.coli"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_SEQ=Not found"
BASE COUNT      221 a    138 c    194 g    216 t
ORIGIN

  Query Match             32.0%;  Score 483.2;  DB 13;  Length 769;
  Best Local Similarity   77.7%;  Pred. No. 2.6e-130;
  Matches  584;  Conservative    0;  Mismatches  168;  Indels    0;  Gaps    0;

Qy        355 GTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAG 414
              ||||| || ||||| || || ||||| || ||||| |||||||||||  | || |||||
Db         18 GTCGGGGTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAA 77

Qy        415 CGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTC 474
               | |||||||| |||||    ||||||   ||||||||    || ||||| ||||||||
```

```
  78 AGAAAATGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTA 137

 475 ACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTT 534
     ||||| ||||| ||||| || |||||||||||||| ||| |||| ||||| || ||  |
 138 ACCGGGGCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTG 197

 535 AAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATG 594
     ||||||||||| ||||| ||||| ||||| |||||||| ||| |  ||||||| || |||
 198 AAGGAAGTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATG 257

 595 GTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTC 654
     |||||||| ||||| |||||||| || |  ||  | |||||||| || ||||| |||||
 258 GTTGATGAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTT 317

 655 GAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGAT 714
     ||||||||||| |  || || || |||||| | || || ||     | ||||||||||||
 318 GAAGATGTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGAT 377

 715 ACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAA 774
     ||||| || || ||||| || ||||||||||| || |||||||| ||| | || ||||||
 378 ACACCTATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAA 437

 775 TTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTAT 834
      | |||||||| ||  |  | || ||||| |||||||| ||||| ||||||||||| |||
 438 CTTGAATGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATAT 497

 835 GGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAG 894
     || ||||| ||||| || ||||| ||||| || ||||| |||||||| || ||||||||
 498 GGGCTTGTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGAT 557

 895 GAACTCATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTC 954
     ||||||||||| || ||||| |||||||||||||||| ||||| || || |||||||| |||
 558 GAACTCATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTC 617

 955 TCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAG 1014
     ||||||||||| |||||| | |||||||||||||| ||  | || || |||||   ||||
 618 TCCAAAGGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAG 677

1015 GGTTACAGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTT 1074
     || ||||  |   | ||||  || ||   ||| |||   |  || |||| | || |  |
 678 GGATACAAGAGCATCATGGGAAACTGTCAAGAAAATGCAAATGTTCTAAAGAAAAGTTTG 737

1075 GAGAAGACAGAAAGGTTCAACATCGTCTCAAA 1106
     ||||| || | | ||||||||||| |||| ||
 738 GAGAAAACCGGACGGTTCAACATCATCTCTAA 769


SULT 14
023643
CUS         BU023643                750 bp    mRNA    linear   EST 23-AUG-2002
FINITION    QHF12C03.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
            clone QHF12C03, mRNA sequence.
CESSION     BU023643
RSION       BU023643.1  GI:22459163
YWORDS      EST.
URCE        Helianthus annuus (common sunflower)
ORGANISM    Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
FERENCE     1  (bases 1 to 750)
AUTHORS     Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
TITLE       Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
JOURNAL     Unpublished
MMENT       Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QHF12  row: C  column: 03.
FEATURES             Location/Qualifiers
     source          1..750
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHF12C03"
                     /lab_host="E.coli"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_SEQ=Not found"
BASE COUNT      221 a    132 c    187 g    210 t
ORIGIN

  Query Match             32.0%;  Score 482.8;  DB 13;  Length 750;
  Best Local Similarity   77.7%;  Pred. No. 3.4e-130;
  Matches  583;  Conservative    0;  Mismatches  167;  Indels    0;  Gaps    0;

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              || ||||| || || ||||| || ||||| |||||||||||  | || |||||  | |||
Db          1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||| |||||    ||||||   ||||||||    || ||||| |||||||| |||||
Db         61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              ||||| ||||| || |||||||||||||| ||| |||| ||||| || ||  | ||||||
Db        121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              ||||| ||||| ||||| ||||| |||||||| ||| |  ||||||| || |||||||||
Db        181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              || ||||| |||||||| || |  ||  | |||||||| || ||||| ||||| ||||||
Db        241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| |  || || || |||||| | || || ||     | |||||||||||||||||
Db        301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || || ||||| || ||||||||||| || |||||||| ||| | || |||||| | |||
Db        361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||| ||  |  | || ||||| |||||||| ||||| ||||||||||| ||||| |||
Db        421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || ||||| || ||||| ||||| || ||||| |||||||| || |||||||| ||||||
Db        481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              ||||| || ||||| |||||||||||||| ||||| || || |||||||| |||||||||
Db        541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
```

```
          ||||| |||||| | |||||||||||||| ||  | || || |||||   |||||| |||
      601 GGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

     1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
          |  |   | ||||| || ||   ||| |||   |  || |||| | ||||  | |||||
      661 AAGAGCATCATGGAAAACTGTCAAGAAAATGCAAATGTTCTAAAGAAAGGTTTGGAGAAA 720

     1081 ACAGAAAGGTTCAACATCGTCTCAAAGGAC 1110
          || | | || |||||||| |||| || |||
      721 ACCGGACGGGTCAACATCATCTCTAAAGAC 750


SULT 15
025264
CUS       BU025264                746 bp    mRNA    linear   EST 23-AUG-2002
FINITION  QHF8J20.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
          clone QHF8J20, mRNA sequence.
CESSION   BU025264
RSION     BU025264.1  GI:22460784
YWORDS    EST.
URCE      Helianthus annuus (common sunflower)
ORGANISM  Helianthus annuus
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          asterids; campanulids; Asterales; Asteraceae; Asteroideae;
          Heliantheae; Helianthus.
FERENCE   1  (bases 1 to 746)
AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
          Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
          ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
          Church,S., Jackson,L. and Bradford,K.
TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
          http://compgenomics.ucdavis.edu/
JOURNAL   Unpublished
MMENT     Contact: Alexander Kozik [R.W.Michelmore]
          Department of Vegetable Crops, R.W.Michelmore Lab
          University of California at Davis (UCD)
          Asmundson Hall, UCD, Davis, CA 95616, USA
          Tel: 1-(530)-742-1742
          Fax: 1-(530)-752-9659
          Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
          belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
          for details.
          Plate: QHF8  row: J  column: 20.
ATURES             Location/Qualifiers
   source          1..746
                   /organism="Helianthus annuus"
                   /mol_type="mRNA"
                   /cultivar="RHA280"
                   /db_xref="taxon:4232"
                   /clone="QHF8J20"
                   /lab_host="E.coli"
                   /clone_lib="QH_EFGHJ sunflower RHA280"
                   /note="Vector: pBRcDNASfiAB; The library was constructed
                   from 11 different sources of RNA from a single genotype.
                   Separate cDNAs were generated using primers that
                   incorporated unique 5' and 3' tags to distinguish each
                   source of RNA. cDNAs were then pooled, size-fractionated,
                   directionally cloned into a custom medium-copy vector and
                   transformations made with four size classes to minimize
                   size bias. Details of each source of RNA and library
                   construction can be obtained at http://cgpdb.ucdavis.edu/
                   TAG_LIB=QH_EFGHJ sunflower RHA280
                   TAG_TISSUE=hulls
                   TAG_SEQ=GCTAGTCGGG"
SE COUNT      219 a    131 c    185 g    211 t
IGIN

Query Match             31.9%;  Score 482;  DB 13;  Length 746;
Best Local Similarity   77.9%;  Pred. No. 5.8e-130;
Matches  581;  Conservative    0;  Mismatches  165;  Indels    0;  Gaps    0;

      361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
```

```
               || ||||| || || ||||| || ||||| |||||||||||  | || |||||  | |||
Db           1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

Qy         421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
               ||||| |||||    ||||||   ||||||||    || ||||| |||||||| |||||
Db          61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

Qy         481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
               ||||| ||||| || ||||||||||||||| ||| |||| ||||| || ||  | ||||||
Db         121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180

Qy         541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
               ||||| ||||| ||||| ||||| |||||||| ||| |  ||||||| || |||||||||
Db         181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

Qy         601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
               || ||||| |||||||| || |  ||  | |||||||| || ||||| ||||| ||||||
Db         241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

Qy         661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
               ||||| |  || || || |||||| | || || ||     | |||||||||||||||||
Db         301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

Qy         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
               || || ||||| || ||||||||||| || |||||||| ||| | || |||||| | |||
Db         361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

Qy         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
               ||||| ||  |  | || |||||| |||||||| ||||| |||||||||||  ||||| |||
Db         421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

Qy         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
               || ||||| || ||||| ||||| || ||||| |||||||| || ||||| |||||||||
Db         481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

Qy         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
               ||||| || |||||| ||||||||||||||| ||||| || || |||||||| |||||||||
Db         541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600

Qy         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
               ||||| |||||| | |||||||||||||| ||  | || || |||||   |||||| |||
Db         601 GGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

Qy        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
               |  |   | ||||| || ||   ||| |||   |  || |||| | ||||  | |||||
Db         661 AAGAGCATCATGGAAAACTGTCAAGAAAATGCAAATGTTCTAAAGAAAGGTTTGGAGAAA 720

Qy        1081 ACAGAAAGGTTCAACATCGTCTCAAA 1106
               || | | ||||||||||| |||| ||
Db         721 ACCGGACGGTTCAACATCATCTCTAA 746

Search completed: October 22, 2003, 13:34:35
Job time : 3001 secs
```

```
                GenCore version 5.1.6
         Copyright (c) 1993 - 2003  Compugen Ltd.

M nucleic - nucleic search, using sw model

un on:          October 22, 2003, 12:44:32 ; Search time 425 Seconds
                                           (without alignments)
                                           9521.971 Million cell updates/sec

itle:           US-10-006-852-1
erfect score:   1509
equence:        1 atggtgctctcccacgccgt..........agacgagtggtatctgctaa 1509

coring table:   IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

earched:        1792395 seqs, 1340900451 residues

otal number of hits satisfying chosen parameters:      3584790

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :        Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                16:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                17:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                %
sult           Query
 No.   Score   Match Length DB  ID                           Description
-------------------------------------------------------------------------------
   1    1509   100.0   1509  14  US-10-006-852-1              Sequence 1, Appli
   2  1507.4    99.9   1509  10  US-09-938-842A-1810          Sequence 1810, Ap
   3   883.4    58.5   1705  14  US-10-006-852-11             Sequence 11, Appl
   4   881.8    58.4   1785  14  US-10-006-852-15             Sequence 15, Appl
   5   870.2    57.7   1665  14  US-10-006-852-3              Sequence 3, Appli
   6   864.2    57.3   1745  14  US-10-005-602-1              Sequence 1, Appli
   7   857.8    56.8   1771  14  US-10-006-852-13             Sequence 13, Appl
   8   854.2    56.6   1482  10  US-09-938-842A-937           Sequence 937, App
   9   834.2    55.3   1479  10  US-09-887-576-820            Sequence 820, App
  10   775.4    51.4   1497  12  US-10-167-547C-1             Sequence 1, Appli
  11     751    49.8   1530  12  US-10-167-547C-3             Sequence 3, Appli
  12     751    49.8   1590  12  US-10-167-547C-5             Sequence 5, Appli
  13   736.6    48.8   1783  14  US-10-006-852-17             Sequence 17, Appl
  14   711.4    47.1   1180  11  US-09-727-758-1              Sequence 1, Appli
  15   709.6    47.0   1455  10  US-09-887-576-782            Sequence 782, App
  16   684.6    45.4   1287  10  US-09-887-576-806            Sequence 806, App
```

```
     17   560.2   37.1   1467  10  US-09-887-576-777          Sequence 777, App
     18   408.6   27.1    894  12  US-10-167-547C-30          Sequence 30, Appl
     19     393   26.0   5160  10  US-09-887-576-443          Sequence 443, App
     20   385.8   25.6   2121  14  US-10-006-852-7            Sequence 7, Appli
     21   360.6   23.9    742  10  US-09-966-881-20           Sequence 20, Appl
     22   351.8   23.3   1946  14  US-10-006-852-9            Sequence 9, Appli
     23   325.4   21.6    807  10  US-09-966-881-21           Sequence 21, Appl
     24     310   20.5   2493  14  US-10-006-852-5            Sequence 5, Appli
     25   283.4   18.8   1410  14  US-10-156-761-4639         Sequence 4639, Ap
 c   26   283.4   18.8 9025608  14  US-10-156-761-1            Sequence 1, Appli
 c   27     278   18.4    493  11  US-09-770-961-610          Sequence 610, App
     28   269.8   17.9   1302  10  US-09-887-576-783          Sequence 783, App
     29   259.4   17.2    440  9  US-09-770-444-922          Sequence 922, App
     30   257.8   17.1   1362  14  US-10-156-761-3588         Sequence 3588, Ap
     31   257.8   17.1 9025608  14  US-10-156-761-1            Sequence 1, Appli
     32   242.6   16.1    387  10  US-09-878-574-1480         Sequence 1480, Ap
     33   187.2   12.4    280  10  US-09-878-574-7302         Sequence 7302, Ap
     34   156.2   10.4    277  10  US-09-878-574-13198        Sequence 13198, A
 c   35   115.2    7.6    477  9  US-09-770-444-28           Sequence 28, Appl
     36   112.4    7.4    781  11  US-09-989-442-73           Sequence 73, Appl
     37      65    4.3    298  12  US-10-167-547C-31          Sequence 31, Appl
 c   38      64    4.2    359  12  US-10-084-843-192          Sequence 192, App
 c   39      64    4.2    359  12  US-10-193-002-187          Sequence 187, App
     40    59.8    4.0   2000  10  US-09-887-576-857          Sequence 857, App
     41    40.8    2.7   1486  9  US-09-443-704-15           Sequence 15, Appl
     42    40.8    2.7   1486  13  US-10-008-118A-15          Sequence 15, Appl
     43      38    2.5 251364  14  US-10-175-523-58           Sequence 58, Appl
     44      38    2.5 251364  14  US-10-175-523-61           Sequence 61, Appl
     45      38    2.5 251364  14  US-10-175-523-79           Sequence 79, Appl

                                ALIGNMENTS

RESULT 1
US-10-006-852-1
; Sequence 1, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1509
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1509)
;   OTHER INFORMATION:
US-10-006-852-1

  Query Match             100.0%;  Score 1509;  DB 14;  Length 1509;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1509;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
b        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

y        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

y        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

y        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

y        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

y        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

/        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

/        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

/        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

/        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

/        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Qy      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

Qy      1501 ATCTGCTAA 1509
             |||||||||
Db      1501 ATCTGCTAA 1509


RESULT 2
US-09-938-842A-1810
; Sequence 1810, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1810
;   LENGTH: 1509
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-1810

  Query Match             99.9%;  Score 1507.4;  DB 10;  Length 1509;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1508;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

Qy        61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

Qy       121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

Qy       181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

Qy       241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
```

```
         301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

         361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

         421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

         481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

         541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

         601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
             |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
         601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

         661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

        1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

        1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

        1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

Qy      1501 ATCTGCTAA 1509
             |||||||||
Db      1501 ATCTGCTAA 1509


RESULT 3
US-10-006-852-11
; Sequence 11, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 1705
;   TYPE: DNA
;   ORGANISM: Nicotiana tabacum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (71)..(1558)
;   OTHER INFORMATION:
US-10-006-852-11

  Query Match             58.5%;  Score 883.4;  DB 14;  Length 1705;
  Best Local Similarity   75.2%;  Pred. No. 1.2e-281;
  Matches 1125;  Conservative    0;  Mismatches  351;  Indels   21;  Gaps    1;

Qy         1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
             ||||| || ||| |  | |  |||||    ||||||||| |||||||||| ||||| ||
Db        71 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 130

Qy        61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
             || || || |||||||| ||||| ||||| |||||||| || || ||||| || ||||||
Db       131 CGATATGTTCGTACTTCTCTTCCGAGGTTTAAGATGCCAGAGAATTCGATACCAAAGGAA 190

Qy       121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
             || || ||||| ||||| || || ||||| ||| | || || |||||| |  | || |||
Db       191 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGACTAAATTTA 250

Qy       181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
             || || |||||||| ||||||||||| || |||||| | ||||| || |||   |||||
Db       251 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAACAAACTGATGATGGATTCCATT 310

Qy       241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
             |||||||| || ||||||||||| || ||||| || ||||| ||||||||||| ||||||
Db       311 AACAAGAATTACGTTGACATGGATGAATACCCTGTAACCACTGAACTTCAGAATCGATGT 370

Qy       301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
             || |||||||| || ||| | || || |||||  | | ||| |  ||||| || || |||
Db       371 GTAAACATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 430

Qy       361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
             || ||||| |||||||| || ||||| || ||| | || || || || |||||| | |||
Db       431 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 490

Qy       421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
             ||||| || ||    |||||  ||||||| |||   || || ||||| |||||||| ||
Db       491 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTGTGACAAGCCCAATATTGTCACTGGT 550
```

```
)y        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              |||||||| || ||||||||||||||||| || ||||| ||||| || || || ||||||
)b        551 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 610

)y        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              || || |||||||| ||||||||||||||||||||| |  ||||||| || ||||| |||
)b        611 GTAAAGTTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 670

)y        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              |||||||| |||||||| || |  ||  | |||||||| || ||||| ||||| ||||||
b         671 GAGAACACAATTTGTGTAGCTGCTATCTTGGGTTCCACACTCAATGGTGAATTTGAAGAT 730

y         661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| | |||||| || ||||||| | || ||||||||||||||||| ||||| || |||
b         731 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 790

y         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || || |||||||| |||||||| ||||| |||||||| ||  | ||||| ||  | |||
b         791 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 850

y         781 TGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||||||||| | || |||||||||||||| || |||||||||||||| ||||| |||
b         851 TGGACTTTAGATTGCCATTGGTGAAGAGTATAAACGTGAGTGGTCACAAATATGGTCTT 910

y         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || || || || ||||||||||  || ||||| || || || || || ||||| ||||| 
b         911 GTTTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTACCTGACGAACTT 970

y         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              |||||||| || ||||||||||||||||| ||||| || || |||||||| ||||| |||
b         971 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1030

y         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              ||||| || ||||| ||||||||||| ||||||||||| || |||||    |||||||||
)        1031 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1090

y        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              |  ||||| |||||||||||   ||| |||   |  || ||||| ||||||||||| || 
)        1091 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1150

'        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              |  | ||| ||||||||  | || || ||    ||||| ||| | || ||||||||  | 
)        1151 AGTGGAAGATTCAACATAATATCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1210

'        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              ||||| | |||    |||| |||||||||||| || || |  ||| | || | |||||||
         1211 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1270

         1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              ||  | || || || || ||||| |||||||| |||||  |||| |||||  | || || 
         1271 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1330

         1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              || ||||||||||||||  | |||||||| ||| |||| || || || || || ||||| 
         1331 ATTAGAGAAGATTTCTCCCGTACACTCGCCGAGCGACTGGTAATAGACATTGAAAAAGTC 1390

         1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
               | |  ||||| ||  | |||||  |||| || |     ||  || |  | |   | | |
         1391 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTGGCCGAGGCG 1450

         1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
              ||  |                      |  | |  |   ||||| |  ||||  ||   |
         1451 AATGG---------------------CAGCGGCGTGCATAAGAAAACAGATAGAGAAGTG 1489

         1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
              |||   || || |  ||||  |||||||| ||||| || || | |||||||| || |
         1490 CAGCTTGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACT 1546
```

SULT 4
-10-006-852-15
Sequence 15, Application US/10006852
Publication No. US20030046732A1

```
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 1785
;   TYPE: DNA
;   ORGANISM: Petunia x hybrida
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (72)..(1571)
;   OTHER INFORMATION:
US-10-006-852-15

  Query Match             58.4%;  Score 881.8;  DB 14;  Length 1785;
  Best Local Similarity   74.9%;  Pred. No. 4.2e-281;
  Matches 1121;  Conservative    0;  Mismatches  367;  Indels    9;  Gaps    1;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||| || ||  |  | || ||| ||   || || ||| | |||||||| || || || 
Db         72 ATGGTTCTATCAAAGACAGTGTCGCAGAGCGATGTGTCCATTCACTCCACGTTTGCTTCT 131

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || || || || ||||| ||||| ||||| || ||||| || || ||||| || || |||
Db        132 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGATAATTCGATACCAAAAGAA 191

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              || || ||||||||||| || || || |||||| | || || || ||| || ||||||| 
Db        192 GCAGCATATCAGATCATAAATGATGAACTGATGTTAGATGGAAACCCAAGGCTGAACTTG 251

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              || || ||||| || ||||||||||| || |||||||||||  | || |||  ||| || 
Db        252 GCTTCTTTTGTTACAACATGGATGGAACCAGAGTGTGATAAGTTGATGATGGACTCTATT 311

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||||||||| ||||| || || || || ||||| || |||||||| ||||||
Db        312 AACAAGAACTATGTTGATATGGATGAATATCCTGTTACCACTGAGCTTCAGAATCGATGT 371

Qy        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              || |||||||| || ||| | || ||||||||  | ||||| |  || || || || |||
Db        372 GTAAACATGATAGCTCATTTGTTTAATGCACCACTTGAAGATGGAGAAACTGCAGTTGGA 431

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              || ||||| |||||||| || || ||||| ||| | || || || || |||||| | |||
Db        432 GTTGGAACTGTTGGATCCTCTGAAGCCATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 491

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              |||||||||||    |||||  |||||||||||   || || ||||||||||| || || 
Db        492 TGGCAGAACAAAATGAAAGCCCAAGGCAAACCCTGTGACAAGCCCAACATTGTTACTGGT 551

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              || ||||| || ||||| ||||||||||| || ||||| ||||| || || || ||||||
Db        552 GCAAATGTCCAGGTGTGCTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 611

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              || ||  | ||||||||||||||||||||||||||| |  ||||||| || ||||| |||
Db        612 GTAAAGCTTAGTGAAGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAGATGGTGGAT 671

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              || |||||||||||||| || |  ||  | |||||||| || ||||||||||| ||||| 
Db        672 GAAAACACCATTTGTGTAGCTGCTATCTTAGGTTCCACCCTCAATGGAGAATTTGAAGAC 731

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| | |||||| |||||||||||||| ||||||||||||||||| ||||| || |||
Db        732 GTTAAGCGCTTGAATGATCTCTTGGTCGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 791
```

```
y         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
b         792 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCGTTCATTTACCCAGAGCTTGAG 851

y         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
b         852 TGGGACTTTAGATTGCCATTAGTGAAGAGCATTAATGTAAGTGGTCACAAATATGGTCTT 911

y         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
b         912 GTCTATGCTGGTATTGGTTGGGTCGTTTGGAGGAACAAGGATGATTTGCCTGATGAACTT 971

y         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
b         972 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTTTCTAAA 1031

y         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
b        1032 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTATGAGGGTTAC 1091

y        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
b        1092 AAGAATGTGATGGAGAATTGTCAAGAAAATGCATCGGTACTAAGAGAAGGGCTAGAAAAG 1151

y        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
b        1152 ACAGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTACCTTTAGTAGCATTCTCTCTT 1211

y        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
b        1212 AAAGACAACAGGCAACACAACGAGTTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1271

y        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
         1272 ATTGTTCCTGCATATACTATGCCACCAAACGCACAACACATTACAGTTCTCAGAGTTGTG 1331

y        1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
         1332 ATCAGAGAAGATTTCTCCCGTACGCTTGCAGAACGACTGGTAAGAGACATCGAAAAAGTC 1391

         1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
         1392 CTTCATGAACTTGACACACTCCCTGCACGTGTCAATGCTAAGCTCGCTGTGGCCGAGGAG 1451

         1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
         1452 CAGGCGGCTGCGAATGGC---------AGCGAGGTGCATAAGAAAACAGATAGCGAAGTG 1502

         1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
         1503 CAGTTGGAGATGATAACTGCATGGAAGAAGTTTGTTGAAGAAAAGAAGAAGAAGACT 1559

SULT 5
-10-006-852-3
Sequence 3, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
 APPLICANT: Kinnersely, Alan M.
 APPLICANT:  Turano, Frank J.
 TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 FILE REFERENCE: 7224-65
 CURRENT APPLICATION NUMBER: US/10/006,852
 CURRENT FILING DATE:  2002-07-01
 PRIOR APPLICATION NUMBER: US 60/246,367
 PRIOR FILING DATE: 2000-11-07
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
  LENGTH: 1665
  TYPE: DNA
  ORGANISM: Arabidopsis thaliana
  FEATURE:
```

```
;   NAME/KEY: CDS
;   LOCATION: (17)..(1498)
;   OTHER INFORMATION:
US-10-006-852-3

  Query Match             57.7%;  Score 870.2;  DB 14;  Length 1665;
  Best Local Similarity   75.1%;  Pred. No. 2.8e-277;
  Matches 1110;  Conservative    0;  Mismatches  348;  Indels   21;  Gaps    1;

Qy         31 GACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCACTTCCTAGGTTC 90
Db         44 GATGAATCTGTCTGCACCATGTTCGGATCTCGCTATGTTCGCACTACACTTCCCAAGTAT 103

Qy         91 AAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTG 150
Db        104 GAGATTGGTGAGAATTCGATACCGAAAGACGCTGCATATCAGATCATAAAAGATGAGCTG 163

Qy        151 ATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCT 210
Db        164 ATGCTTGATGGTAACCCGAGGCTTAACCTAGCTTCGTTTGTGACTACATGGATGGAACCA 223

Qy        211 GAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTAC 270
Db        224 GAGTGTGACAAACTCATCATGGACTCTATCAACAAGAACTACGTTGATATGGATGAGTAC 283

Qy        271 CCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATCTATTCAATGCA 330
Db        284 CCTGTCACAACTGAGCTCCAGAACCGATGTGTAAACATTATAGCTCGACTGTTCAATGCG 343

Qy        331 CCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATA 390
Db        344 CCACTCGAGGAATCTGAGACGGCGGTGGGAGTAGGGACAGTTGGTTCTTCAGAAGCCATC 403

Qy        391 ATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAA 450
Db        404 ATGTTAGCCGGATTGGCCTTCAAAAGAAAATGGCAGAACAAACGCAAGGCTGAGGGTAAA 463

Qy        451 CCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTC 510
Db        464 CCCTATGACAAACCCAACATTGTCACTGGAGCCAATGTTCAAGTTTGCTGGGAGAAATTC 523

Qy        511 GCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATG 570
Db        524 GCTCGGTACTTCGAGGTGGAGCTAAAGGAAGTAAACCTAAGTGAAGGTTACTACGTGATG 583

Qy        571 GACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTT 630
Db        584 GATCCAGACAAAGCAGCAGAAATGGTAGACGAGAACACAATCTGTGTCGCAGCCATATTG 643

Qy        631 GGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAA 690
Db        644 GGATCCACACTCAACGGTGAGTTCGAAGACGTGAAACGTCTCAATGACTTGCTAGTCAAG 703

Qy        691 AAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTC 750
Db        704 AAAAACGAGGAGACTGGTTGGAACACACCGATCCACGTGGATGCAGCAAGTGGAGGGTTC 763

Qy        751 ATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGT 810
Db        764 ATAGCTCCGTTTATCTATCCTGAATTAGAATGGGACTTTAGACTTCCTTTGGTTAAGAGT 823

Qy        811 ATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGG 870
Db        824 ATCAACGTGAGTGGTCACAAGTATGGACTGGTCTATGCTGGTATTGGTTGGGTCGTGTGG 883

Qy        871 AGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGAC 930
Db        884 AGGGCAGCAGAGGATTTGCCTGAAGAGCTTATCTTTCATATTAATTATCTTGGTGCTGAT 943

Qy        931 CAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTAC 990
Db        944 CAACCCACTTTCACTCTCAATTTCTCCAAGGGATCGAGCCAAATTATTGCTCAATACTAC 1003
```

```
  991 CAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAAT 1050
      || || || ||  | ||   |||||| |||| ||||||||||||||||||||||| ||||||
 1004 CAGCTCATTCGTCTTGGATTCGAGGGGTACAAAAATGTGATGGAGAATTGCATAGAGAAC 1063

 1051 ATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGAC 1110
      ||| | || || |  |||||  | ||||| |||||  | |||||||| ||||||||||||
 1064 ATGGTGGTTCTCAAAGAAGGGATAGAGAAAACAGAGCGTTTCAACATAGTCTCAAAGGAC 1123

 1111 GAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAA 1170
       | ||||||||| | || || |||||  | || ||    || |  ||||  ||||||||
 1124 CAAGGAGTGCCAGTCGTAGCCTTCTCTCTCAAGGACCATAGTTTCCACAACGAGTTCGAG 1183

 1171 ATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAAT 1230
      ||||| || ||||| ||  | | ||| |||||| || || || ||||| |||||| |  ||
 1184 ATCTCTGAGATGCTACGTCGTTTTGGCTGGATCGTCCCAGCTTACACTATGCCTGCCGAT 1243

 1231 GCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCT 1290
      ||||| ||||||||| ||||| ||||| || ||||| ||||||||||| |||||||||||
 1244 GCACAGCACATCACGGTTCTGCGTGTTGTCATCAGGGAAGATTTCTCAAGAACACTCGCG 1303

 1291 GAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGCTTCCTTCGAGA 1350
      |||||||||||    |||||   ||| ||| | | ||||||| |||    | ||||| |
 1304 GAGAGACTTGTTGCTGATATTTCGAAGGTGCTTCATGAGCTAGATACCTTGCCTTCCAAG 1363

 1351 GTGATTCACAAAATATCACTTGGACAAGAGAAGAGTGAATCTAACAGCGATAACTTGATG 1410
       |                     |    ||||||  | |    |  |          |
 1364 AT---------------------ATCTAAGAAGATGGGAATAGAAGGGATCGCGGAAAAT 1402

 1411 GTCACGGTGAAGAAGAGCGATATCGACAAGCAGAGAGATATCATCACTGGCTGGAAGAAG 1470
      || | || ||||||||  || |  || || |  | ||  || ||   ||| |||| ||||
 1403 GTAAAGGAGAAGAAGATGGAGAAGGAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAG 1462

 1471 TTTGTCGCCGACAGGAAGAAGACGAGTGGTATCTGCTAA 1509
      |||||    || |||||||||| || |||| | ||||||
 1463 TTTGTGAAGGAGAGGAAGAAGATGAATGGTGTGTGCTAA 1501
```

SULT 6
-10-005-602-1
Sequence 1, Application US/10005602
Publication No. US20030110530A1
GENERAL INFORMATION:
 APPLICANT: Barry Shelp
 APPLICANT: Alan Bown
 TITLE OF INVENTION: TRANSGENIC PLANTS HAVING REDUCED
 TITLE OF INVENTION: SUSCEPTIBILITY TO INVERTEBRATE PESTS
 FILE REFERENCE: P84US3
 CURRENT APPLICATION NUMBER: US/10/005,602
 CURRENT FILING DATE: 2002-12-31
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 1
 LENGTH: 1745
 TYPE: DNA
 ORGANISM: tobacco plant
-10-005-602-1

```
Query Match             57.3%;  Score 864.2;  DB 14;  Length 1745;
Best Local Similarity   74.3%;  Pred. No. 2.8e-275;
Matches 1113;  Conservative    0;  Mismatches  363;  Indels   21;  Gaps    1;

    1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
      ||||| || ||| |  | |  |||||    ||||||||| |||||||||| ||||| ||
   31 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 90

   61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
      || || || || ||||| ||||| ||||| |||||||| || || || || || ||||||
   91 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAGATGCCAGAGAATTCAATACCAAAGGAA 150

  121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
      || || |||||||| || ||  || ||||| ||| | || || |||||| || | || |||
  151 GCAGCATATCAGATTATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 210
```

```
Qy       181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
             || || || || || ||||||||||||||| || ||| |||   | || |||   |||||
Db       211 GCATCTTTCGTTACAACATGGATGGAGCCAGAATGTAATACGTTAATGATGGATTCCATT 270

Qy       241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
             ||||||||||| ||||||||||| || |||||| || |||| || |||||||| ||||||
Db       271 AACAAGAACTACGTTGACATGGATGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 330

Qy       301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
             || || ||||| || ||| | || ||||||||  | | ||| |  ||||| || || |||
Db       331 GTAAATATGATAGCTCATTTGTTTAATGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 390

Qy       361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
             || ||||| |||||||| ||  || || || ||| | || || || |||||  ||  | |||
Db       391 GTTGGAACTGTTGGATCCTCTGAAGCTATTATGCTTGCTGGATTAGCCTTTAAAAGAAAA 450

Qy       421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
             ||||| || ||    |||||  ||||||| ||| | |||| ||||| || ||||||||
Db       451 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTTTGATAAGCCCAATATCGTCACCGGT 510

Qy       481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
             || ||||| || ||||||||||||||||||| || ||||| |||||| || ||   | || |||
Db       511 GCTAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 570

Qy       541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
             || ||||||||||| |||||||||||||||||||||| |  ||||||| || ||||| |||
Db       571 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 630

Qy       601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
             ||||| |||||||||||||| |  ||  | ||||| || || ||||| || ||||| ||||||
Db       631 GAGAATACCATTTGTGTTGCTGCTATCTTAGGTTCAACACTCAATGGTGAATTTGAAGAT 690

Qy       661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
             ||||| |  ||||| || || ||| | || |||||||||||||||||| ||||| || |||
Db       691 GTTAAGCGTTTGAATGACCTTTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 750

Qy       721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
             || || |||||||| |||||||| ||||| |||||||| ||  | ||||| ||  | |||
Db       751 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 810

Qy       781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
             ||||||||||||| | || ||||||||||||||| ||||||||||||||||||| ||||| |||
Db       811 TGGGACTTTAGATTGCCATTGGTGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 870

Qy       841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
             || || || || ||||||||||  || ||||| || || || || |||||||| |||||
Db       871 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 930

Qy       901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
             || ||||| |||||||| |||||||||| ||||| || || |||||||| ||||| |||
Db       931 ATTTTCCACATCAATTACCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 990

Qy       961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
             ||||| || ||||| |||||||||| |||||||||||| || |||||    |||||||||
Db       991 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1050

Qy      1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
             |  ||||| |||||||||||   ||| |||   |  ||  |||| |||||| |||| ||
Db      1051 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTATTAAGAGAAGGAATTGAAAAA 1110

Qy      1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
             |  | ||| ||||||||  |||| || ||    ||||| ||  | || || || ||  |
Db      1111 AGTGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTTCCCTTAGTAGCATTTTCTCTT 1170

Qy      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
             ||||| | |||    |||| ||||||||||||| || || |  ||| | || | |||||||
Db      1171 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1230

Qy      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
             || || || || || || ||| |||||||| |||||  | || |||||  | || ||
Db      1231 ATTGTTCCTGCATATACTATGCCACCAAATGCTCAACATGTTACAGTTCTCAGAGTTGTC 1290

Qy      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
```

```
     || ||||||||||||||  | ||||| || ||| |||| || || || || || |||||
1291 ATTAGAGAAGATTTCTCCCGCACACTAGCGGAGCGACTGGTAATAGACATTGAAAAAGTC 1350

1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
      | |  ||||| ||    |||||  |||| || |     ||  || |  | |   | | |
1351 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTGGCCGAGGCG 1410

1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
     ||  |                      |  | |  |   ||||| |  ||||  ||   |
1411 AATGG---------------------CAGCGGCGTGCATAAGAAAACAGATAGAGAAGTG 1449

1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
     |||  ||| || |  ||||  |||||||| ||||| || || | |||||||| || |
1450 CAGCTAGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACT 1506

SULT 7
-10-006-852-13
Sequence 13, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
 APPLICANT: Kinnersely, Alan M.
 APPLICANT:  Turano, Frank J.
 TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 FILE REFERENCE: 7224-65
 CURRENT APPLICATION NUMBER: US/10/006,852
 CURRENT FILING DATE:  2002-07-01
 PRIOR APPLICATION NUMBER: US 60/246,367
 PRIOR FILING DATE: 2000-11-07
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
  LENGTH: 1771
  TYPE: DNA
  ORGANISM: Nicotiana tabacum
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (67)..(1554)
  OTHER INFORMATION:
-10-006-852-13

Query Match             56.8%;  Score 857.8;  DB 14;  Length 1771;
Best Local Similarity   74.1%;  Pred. No. 3.8e-273;
Matches 1109;  Conservative    0;  Mismatches  367;  Indels   21;  Gaps    1;

   1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
     ||||| || ||| |  | |  |||||    ||||||||||| |||||||| ||||| ||
  67 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCGTTCACTCCACTTTCGCCTCC 126

  61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
     || || || || ||||| ||||| ||||| || ||||| || || || || || ||||||
 127 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGAGAATTCAATACCAAAGGAA 186

 121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
     || || |||||||| || || || ||||| ||| | || || |||||| || | || |||
 187 GCAGCATATCAGATTATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 246

 181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
     || || || || || |||||||||||||| || ||| |||   | || |||   |||||
 247 GCATCTTTCGTTACAACATGGATGGAGCCAGAATGTAATACGTTAATGATGGATTCCATT 306

 241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
     ||||||||||| ||||||||||| || ||||| || ||||| || |||||||| ||||||
 307 AACAAGAACTACGTTGACATGGATGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 366

 301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
     || || ||||| || ||| | || ||||||||  | | ||| |  ||||| || || |||
 367 GTAAATATGATAGCTCATTTGTTTAATGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 426

 361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
     || ||||| |||||||| || || || || ||| | || || || ||||| ||| | |||
 427 GTTGGAACTGTTGGATCCTCTGAAGCTATTATGCTTGCTGGATTAGCCTTTAAGAGAAAA 486

Qy   421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
         ||||| || ||    |||||  ||||||| ||| | ||||| ||||| |||||||||||
Db   487 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTTTGATAAGCCCAATATTGTCACCGGT 546

Qy   481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
         || ||||| || ||||||||||||||||| || ||||| ||||| || ||  | || |||
Db   547 GCTAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 606

Qy   541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
         || ||||||||||| ||||||||||||||||||||| |  ||||||| || ||||| |||
Db   607 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 666

Qy   601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
         ||||| |||||||||||||| |  ||  | ||||| || || ||||| ||||| ||||||
Db   667 GAGAATACCATTTGTGTTGCTGCTATCTTAGGTTCAACACTCAATGGTGAATTTGAAGAT 726

Qy   661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
         ||||| |  ||||| || || ||| | || ||||||||||||||||| ||||| || |||
Db   727 GTTAAGCGTTTGAATGACCTTTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 786

Qy   721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
         || || |||||||| |||||||| ||||| |||||||| ||  | ||||| ||  | |||
Db   787 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 846

Qy   781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
         |||||||||||| | || |||| ||||||||| ||||||||||||||||||||||| |||
Db   847 TGGGACTTTAGATTGCCATTGGAGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 906

Qy   841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
         || || || || ||||||||||  || ||||| || || || || |||||||| |||||
Db   907 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 966

Qy   901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
         || ||||| |||||||| ||||||||||| ||||| || || |||||||| ||||| |||
Db   967 ATTTTCCACATCAATTACCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1026

Qy   961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
         ||||| || ||||| ||||||||||| ||||||||||| || |||||    |||||||||
Db  1027 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1086

Qy  1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
         |  ||||| |||||||||||   ||| |||   |  ||  |||| |||||| |||| ||
Db  1087 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTATTAAGAGAAGGAATTGAAAAA 1146

Qy  1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
         |  | ||| ||||||||  |||| || ||    ||||| ||  | || || || ||  |
Db  1147 AGTGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTTCCCTTAGTAGCATTTTCTCTT 1206

Qy  1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
         ||||| | |||    |||| |||||||||||| || || |  ||| | || | |||||||
Db  1207 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1266

Qy  1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
         || || | ||| || || ||||| |||||||| |||||  |||| |||||  | || ||
Db  1267 ATTGTTCTGGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1326

Qy  1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
         || ||||||||||||||  | ||||| || ||| |||| || || || || || |||||
Db  1327 ATTAGAGAAGATTTCTCCCGCACACTAGCGGAGCGACTGGTAATAGACATTGAAAAAGTC 1386

Qy  1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
          | |  |   | ||    |||||  |||| || |     ||  || |  | |   | | |
Db  1387 TTCCACGGAGTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTGGCCGAGGCG 1446

Qy  1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
         ||  |                      |  | |  |   ||||| |  ||||  ||   |
Db  1447 AATGG---------------------CAGCGGCGTGCATAAGAAAACAGATAGAGAAGTG 1485

Qy  1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
         |||  ||| || |  ||||  |||  ||| ||||| || || | |||||||| || |
Db  1486 CAGCTAGAGATTACTACTGCATGGTTGAAATTTGTTGCTGATAAGAAGAAGAAGACT 1542
```

```
SULT 8
-09-938-842A-937
Sequence 937, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
 APPLICANT: Harper, Jeff
 APPLICANT:  Kreps, Joel
 APPLICANT:  Wang, Xun
 APPLICANT:  Zhu, Tong
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 TITLE OF INVENTION:  SAME, AND METHODS OF USE
 FILE REFERENCE: SCRIP1300-3
 CURRENT APPLICATION NUMBER: US/09/938,842A
 CURRENT FILING DATE:  2001-08-24
 PRIOR APPLICATION NUMBER: US 60/227,866
 PRIOR FILING DATE: 2000-08-24
 PRIOR APPLICATION NUMBER: US 60/264,647
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: US 60/300,111
 PRIOR FILING DATE: 2001-06-22
 NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 937
  LENGTH: 1482
  TYPE: DNA
  ORGANISM: Arabidopsis thaliana
-09-938-842A-937

Query Match            56.6%;  Score 854.2;  DB 10;  Length 1482;
Best Local Similarity  76.0%;  Pred. No. 5.3e-272;
Matches 1054;  Conservative    0;  Mismatches  333;  Indels    0;  Gaps    0;

    1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
      |||||  | ||  |  | || || || || |||||  |||| || || || || || 
    1 ATGGTTTTGTCTAAGACAGTTTCCGAATCTGATGTCTCAATCCATTCAACTTTTGCTTCT 60

   61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
      ||||||||||| |  || ||||||  | ||| | ||||| || ||||| || || || |||
   61 CGTTACGTCCGCAACTCTCTTCCACGATTCGAAATGCCTGAGAACTCAATCCCAAAAGAA 120

  121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
      || || || || |||||||||||||||||| ||||| || |  | || ||| | ||||| ||
  121 GCAGCTTACCAAATCATCAACGACGAGCTAATGCTCGATGGTAACCCAAGGCTGAACCTA 180

  181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
      || ||||| ||||| ||||||||||||||| || ||||| || ||||| |||   ||||||
  181 GCTTCCTTCGTGACCACATGGATGGAGCCAGAATGTGACAAGCTCATGATGGAGTCCATC 240

  241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
      ||||||||||| || |||||||||||||||||| ||||||||| || ||||||||||||||||
  241 AACAAGAACTACGTCGACATGGACGAGTACCCTGTCACCACTGAGCTTCAGAACCGATGT 300

  301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
      || ||||||||| |||| ||| ||||| || ||| | |  || |  ||  | ||||| || 
  301 GTTAACATGATAGCACGTCTCTTCAACGCGCCGCTTGGTGACGGTGAAGCTGCCGTTGGT 360

  361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
      || || ||||| ||||| |||||||| || ||||||||||||||||| || ||| |  ||
  361 GTTGGCACCGTCGGATCGTCGGAGGCGATTATGTTGGCCGGTTTGGCTTTTAAGAGACAA 420

  421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
      |||||||| ||||| || ||  ||||    ||    ||||| ||||| || || ||||| 
  421 TGGCAGAATAAGCGTAAGGCCCAAGGGCTTCCTTATGATAAGCCCAATATCGTAACCGGT 480

  481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
      || ||||| || || || |||||||||||||| ||||| || || || || |||||||||
  481 GCTAATGTCCAGGTTTGCTGGGAGAAATTCGCAAGGTATTTCGAAGTGGAGCTTAAGGAA 540

  541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
      |||||  | || ||||  || || ||||||||||||  | | || || || ||||| || 
  541 GTGAACCTAAGAGAAGACTATTACGTGATGGACCCTGTAAAGGCGGTCGAAATGGTAGAC 600

  601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
      || ||||| ||||||||| || | ||| || ||||| ||  | |  || |||||||||||
Db        601 GAAAACACAATTTGTGTCGCTGCCATCCTCGGTTCAACGTTAACCGGTGAATTCGAAGAC 660

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| ||| | ||||| ||| | ||||| || |||||  ||||||||||||||| || |||
Db        661 GTTAAGCTCCTCAACGACCTCCTTGTCGAGAAAAACAAGCAAACCGGATGGGACACGCCA 720

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || |||||||| || || |||| ||  || |||||| ||||| |||||||||||  |||| 
Db        721 ATACACGTGGACGCAGCGAGTGGTGGGTTTATTGCTCCGTTCTTGTATCCGGAGCTGGAG 780

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||| ||  | || || || ||| |||||||| ||||||||||||||||||| || ||  | 
Db        781 TGGGATTTCCGGCTACCGTTGGTTAAGAGTATTAATGTGAGTGGTCACAAATACGGTTTG 840

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || ||||| || |||||||||||  | ||||||| ||||   |||||||||||| ||||| 
Db        841 GTTTACGCCGGTATTGGTTGGGTTGTATGGAGAACCAAAACCGATTTGCCTGATGAACTT 900

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              ||||||||||||||||||||||| ||||| ||||| ||||||||| ||||| |||||||||
Db        901 ATCTTCCATATCAATTATCTTGGCGCTGATCAACCAACCTTTACACTCAACTTCTCCAAA 960

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              |||||||||||||| ||||||||| |||||||||| || || ||  | ||   ||||||||| 
Db        961 GGTTCAAGTCAAGTGATTGCTCAGTACTACCAGCTGATTCGTCTTGGATTCGAGGGTTAT 1020

Qy       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
               | ||||||||||||| |||||  | || ||||||| || |||||  ||||| | ||||| 
Db       1021 CGCAATGTGATGGATAATTGTCGGGAAAACATGATGGTACTAAGACAAGGATTAGAGAAA 1080

Qy       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              || | | | || || |||||||| || ||  | || || ||  | || || || ||  | 
Db       1081 ACGGGACGTTTTAAAATCGTCTCCAAAGAAAACGGTGTTCCGTTAGTGGCGTTTTCTCTC 1140

Qy       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              |||||||| ||| | ||||  ||||||||   |  || | |  || ||  |  |  || |||
Db       1141 AAAGATAGTAGCCGCCACAACGAGTTCGAGGTGGCCCATACACTCCGTCGCTTCGGCTGG 1200

Qy       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              || || ||||||||||| ||||||| |  |||| || ||  ||||||| ||||||| || |||
Db       1201 ATCGTTCCGGCCTACACGATGCCTGCGGATGCGCAGCATGTCACTGTCCTTCGAGTTGTT 1260

Qy       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              ||| ||||||||||||||  ||||  | || |||||| | ||    ||| | ||||| || 
Db       1261 ATCCGAGAAGATTTCTCTCGAACCTTAGCCGAGAGATTGGTAGCTGATTTCGAGAAGGTT 1320

Qy       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
               | |  ||||||||||  ||||||  |||| ||      ||| ||  |   |||| |||  
Db       1321 CTACACGAGCTCGATACGCTTCCGGCGAGGGTTCACGCCAAGATGGCTAATGGAAAAGTT 1380

Qy       1381 AAGAGTG 1387
              ||  |||
Db       1381 AACGGTG 1387


RESULT 9
US-09-887-576-820
; Sequence 820, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
;  APPLICANT: Budworth, P.
;  APPLICANT:  Brown, D.
;  APPLICANT:  Chang, H.
;  APPLICANT:  Zhu, T.
;  APPLICANT:  Han, B.
;  APPLICANT:  Wang, X.
;  APPLICANT:  Cooper, Bret
;  TITLE OF INVENTION: Promoters for regulation of plant expression
;  FILE REFERENCE: 1360.001US1
;  CURRENT APPLICATION NUMBER: US/09/887,576
;  CURRENT FILING DATE:  2001-06-25
;  PRIOR APPLICATION NUMBER: US 60/213,848
;  PRIOR FILING DATE: 2000-06-23
```

```
 PRIOR APPLICATION NUMBER: US 60/214,087
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US 60/258,692
 PRIOR FILING DATE: 2000-12-29
 NUMBER OF SEQ ID NOS: 875
 SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 820
  LENGTH: 1479
  TYPE: DNA
  ORGANISM: Oryza sativa
-09-887-576-820

Query Match             55.3%;  Score 834.2;  DB 10;  Length 1479;
Best Local Similarity   75.9%;  Pred. No. 2.3e-265;
Matches 1031;  Conservative    0;  Mismatches  328;  Indels    0;  Gaps    0;

            1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              |||||||||||| | ||||| || |||   ||| | ||||| |||||||| ||||| ||
            1 ATGGTGCTCTCCAAGGCCGTCTCCGAGAGTGACATGTCCGTGCACTCCACCTTCGCCTCC 60

           61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || ||||||||  | || || || |||| |  ||||||||| |||||||| || |||||
           61 CGCTACGTCCGCGCCTCCCTCCCAAGGTACCGGATGCCGGAGAACTCGATCCCGAAGGAG 120

          121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              |||||||| |||||||||||||||||||||||||| ||||| || || ||| ||||| |
          121 GCGGCGTACCAGATCATCAACGACGAGCTGATGCTGGACGGCAACCCGCGGCTGAACCTG 180

          181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              || || || || || || ||||||||||| ||||| || || ||||||||| || |||||
          181 GCGTCGTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATCATGGCCGCCATC 240

          241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||| || ||||||||||||||||||||||||||||| || |||||||| ||
          241 AACAAGAACTACGTCGACATGGACGAGTACCCCGTCACCACCGAGCTCCAGAACCGGTGC 300

          301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              ||||||||||| || || || ||| | || ||| | |  ||||  ||||| || || ||
          301 GTGAACATGATCGCGCACCTGTTCCACGCGCCGCTCGGGGAGGACGAGACGGCGGTGGGC 360

          361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              || || || || || || ||||||||||| ||| |||||||  |||||||||||||
          361 GTGGGCACGGTGGGTTCGTCGGAGGCCATCATGCTGGCCGGGCTGGCCTTCAAGCGGCGG 420

          421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||||||||||||||| || || || || ||  |||| || ||||||||  |||||||
          421 TGGCAGAACAAGCGCAAGGCCGAGGGGAAGCCGTTCGACAAGCCCAACATCATCACCGGC 480

          481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              ||||| || || ||||| |||||||| |||||  | |||| ||||| || || |||||
          481 GCCAACGTGCAGGTGTGCTGGGAGAAGTTCGCCCGCTACTTCGAGGTGGAGCTCAAGGAG 540

          541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              |||||  |  | || || ||||| || ||||||||  |  | || || || |||||  |
          541 GTGAAGCTCCGCGACGGCTACTACGTCATGGACCCCGAGAAGGCCGTCGACATGGTCAAC 600

          601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              ||||||||||| || || || | ||| || || ||||| || || || || ||||| ||
          601 GAGAACACCATCTGCGTCGCCGCCATCCTCGGCTCCACCCTCAACGGCGAGTTCGAGGAC 660

y         661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              || || ||  | ||||| ||| | | | | |||||||| || || || ||||| || ||
b         661 GTCAAGCTACTCAACGACCTCCTCGACAAGAAGAACAAGGAGACTGGGTGGGAGACGCCG 720

y         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              ||||||||||| ||||| ||  || || ||||| || |||||  |||| ||||||  ||||
b         721 ATCCACGTGGACGCGGCGAGCGGCGGGTTCATCGCGCCGTTCCTGTACCCGGAGCTGGAG 780

y         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||||||  | || || | ||||||||| ||||| ||||| |||||||||||| || ||
b         781 TGGGACTTCCGGCTGCCGTGGGTGAAGAGCATCAACGTGAGCGGTCACAAGTACGGGCTC 840

y         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || ||||| || || || |||   |||||| | ||||| |||||  |||| ||||| |||
Db        841 GTCTACGCCGGCATCGGCTGGTGCATCTGGCGCAACAAGGAGGACCTGCCCGAGGAGCTC 900

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              |||||||| ||||| || || ||  | ||||| || ||||| || ||||| ||||||||
Db        901 ATCTTCCACATCAACTACCTCGGCACCGACCAGCCAACCTTCACCCTCAACTTCTCCAAG 960

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              || || || || ||||| || | ||||||||| || |||||    |||  |||||| |||
Db        961 GGCTCCAGCCAGGTCATCGCCCAGTACTACCAGCTCATCCGCCACGGCTTCGAGGGGTAC 1020

Qy       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              || ||  | |||||||| |||   |||||   ||| || || | ||||||| || | ||||
Db       1021 AGGAACATCATGGAGAACTGCCACGAGAACGCGATGGTGCTGAAGGAAGGGCTGGTGAAG 1080

Qy       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              || |  |||||| ||||||| || |||||||| || |||||| || || || |||||  |
Db       1081 ACGGGGAGGTTCGACATCGTGTCCAAGGACGAAGGGGTGCCGCTGGTGGCGTTCTCGCTC 1140

Qy       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              || ||  | ||| | |||   |||||||| ||||||||||||||| ||| | |  || |||
Db       1141 AAGGACCGGAGCCGGCACGACGAGTTCGAGATCTCCGACATGCTGCGCCGCTTCGGCTGG 1200

Qy       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              || |||||||| ||||| ||||| ||  | || || ||| |||| || || || |||||
Db       1201 ATCGTGCCGGCGTACACCATGCCGCCCGACGCCCAGCACGTCACGGTGCTCCGCGTGGTC 1260

Qy       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              ||| | || || |||    | || ||||| ||| | || ||  |||| || ||||| |||
Db       1261 ATCCGGGAGGAGTTCAGCCGCACCCTCGCCGAGCGCCTCGTCCTCGACATCGAGAAGGTG 1320

Qy       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCAC 1359
              |||    ||||||| | ||| || || ||  | || | |
Db       1321 ATGTACCAGCTCGACGCGCTCCCCTCCAGGCTCATGCCC 1359


RESULT 10
US-10-167-547C-1
; Sequence 1, Application US/10167547C
; Publication No. US20030170653A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  APPLICANT:  Damude, Howard G.
;  TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
;  TITLE OF INVENTION:  Butyrolactone and its Intermediates
;  FILE REFERENCE: CL1804 US NA
;  CURRENT APPLICATION NUMBER: US/10/167,547C
;  CURRENT FILING DATE:  2003-03-17
;  PRIOR APPLICATION NUMBER: 60/297198
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: Microsoft Office 07
; SEQ ID NO 1
;   LENGTH: 1497
;   TYPE: DNA
;   ORGANISM: Alstroemeria
US-10-167-547C-1

  Query Match             51.4%;  Score 775.4;  DB 12;  Length 1497;
  Best Local Similarity   70.6%;  Pred. No. 7.5e-246;
  Matches 1066;  Conservative    0;  Mismatches  431;  Indels   12;  Gaps    2;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||| ||||||  |||||| || ||      || | | |||||||| ||| ||||| ||
Db          1 ATGGTTCTCTCCAGCGCCGTCTCCGACAACACCGGCCCAGTCCACTGCACCTTCGCGTCG 60

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || ||||||||   | | | | |  |||||| ||||||||| || ||||| || |||||
Db         61 CGCTACGTCCGCGATGCGCCTGCCCGGTTCAGGATGCCGGAGAATTCGATACCCAAGGAC 120

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
               ||||||| |||||| |||||||||||||||||||| ||||||||| ||  || ||||| |
Db        121 ACGGCGTACCAGATCGTCAACGACGAGCTGATGCTGGACGGGAACCCCAGGCTGAACCTG 180
```

```
       181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
           || || || || ||||| ||||||||||| ||||| |||   ||||| ||  || ||  |
       181 GCGTCGTTCGTCACGACCTGGATGGAGCCGGAGTGCGATCGCCTCATGATCGCCGCCGCC 240

       241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
           |||||||||||||| |||||||||||||||||| ||||| ||||||| || ||||| |||
       241 AACAAGAACTATGTCGACATGGACGAGTACCCGGTCACAACCGAGCTACAGAATCGCTGT 300

       301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
           || || || || || || || |||||||| ||  | |  || | ||| || || || |||
       301 GTTAATATAATAGCCCACCTCTTCAATGCGCCTATTGGGGATGAGGAAACAGCAGTAGGA 360

       361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
           || || || || || |||||| || || |||||| | || || ||||| ||||||| | || 
       361 GTTGGGACAGTGGGGTCATCAGAAGCAATAATGCTTGCAGGCTTGGCATTCAAGAGAAAG 420

       421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
           |||||||||||    |||||| || || || ||    || || ||||| |||||||||| 
       421 TGGCAGAACAAAAAAAAAGCAGAGGGGAAGCCATATGACAAGCCCAATATTGTCACCGGT 480

       481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
           || || ||||| || || |||||||||||||||||||| ||||| |||||||||| || |||
       481 GCAAACGTTCAGGTTTGCTGGGAGAAATTCGCTAGGTATTTTGAAGTTGAACTGAAAGAA 540

       541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
           |||||  |||| || || || ||  | ||||||||  |  | ||||| || ||||| |||
)      541 GTGAAGCTGAGGGAGGGTTATTACATCATGGACCCAGAGAAGGCTGTGGAAATGGTGGAT 600

       601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
           ||||| ||||| |||||||| |  ||  | || || || |||| |||||| |||||||||
       601 GAGAATACCATATGTGTTGCTGCTATCTTGGGCTCAACCCTTACTGGAGAGTTCGAAGAT 660

       661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
           ||||||||  ||||| | ||  | |  || |||||||| ||||| || ||||| ||||| 
       661 GTTAAACTACTGAACAAACTTCTTGAAGAGAAGAACAAGGAAACTGGGTGGGACACACCC 720

       721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
           || || || ||||| || ||||| |||||||||||| || ||| | || || ||| |||||
       721 ATTCATGTTGATGCTGCTAGTGGTGGATTCATTGCTCCTTTTCTATACCCAGAACTGGAA 780

y      781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
           ||||| ||  || | ||  ||||||||||||| ||||| || || ||||| ||||| |||
b      781 TGGGATTTCCGATTACCACTGGTGAAGAGTATTAATGTCAGCGGACACAAATATGGCCTT 840

y      841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
           || || |||||  | ||||||||  ||||||| |||||||| ||| | ||||| || |||
b      841 GTTTATGCAGGTGTGGGTTGGGTTGTCTGGAGGAACAAAGAAGATCTTCCTGAAGAGCTC 900

y      901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
           || |||||||| || || ||||| || || || || || || || ||||||||||| |||
b      901 ATTTTCCATATAAACTACCTTGGGGCAGATCAGCCTACTTTCACCCTCAATTTCTCTAAA 960

y      961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
           |||||||| ||  | |||||||||||| || ||| | || ||  | |      |||| || 
b      961 GGTTCAAGCCAGATAATTGCTCAATATTATCAATTCATTCGCCTTGGTTTTCAGGGGTAT 1020

y     1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
           |  ||  | ||||| || ||||  ||||| |  |   | || || ||||| ||  || ||
b     1021 AAGAACATAATGGAAAACTGCATGGAGAACACAAGAATACTGAGAGAAGGTCTGCAGGAG 1080

y     1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
           || |   | ||| | || ||||| || ||    || ||||| ||||| || ||  |  | 
Db    1081 ACGGGCCGTTTCGAGATAGTCTCCAAAGATATTGGGGTGCCTCTTGTTGCATTTGCTCTC 1140

Qy    1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
           || || ||||||    ||||||  || || ||  | |||    |  | ||||  ||||||
Db    1141 AAGGACAGCAGCCAGTACACTGTCTTTGAGATAGCGGACGCCATGAGAAGGTTCGGATGG 1200

Qy    1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
           ||  | || || ||||| |||||     | ||  | |||||  |||| || |||||||||
Db    1201 ATCATTCCTGCATACACCATGCCAAAGGACGCGGAGCACATAGCTGTCCTCCGTGTGGTT 1260
```

```
Qy    1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
           ||||| || ||||||   || |  || |||||| | || || |  || ||  |||| |||
Db    1261 ATCAGGGAGGATTTCAGCAGGAGCCTTGCTGAGCGCCTAGTTAATGACATGAAGAAGGTG 1320

Qy    1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
            ||  |||||| || |  |||||          | || ||  |   |  |||  || |  
Db    1321 CTGGCTGAGCTGGACGTACTTCC---------CAGTCGCATCACCACCATTGCCCATGTT 1371

Qy    1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
           | |  ||      | | ||||||       |    | | ||||||||     | || |  
Db    1372 ACGGCTG---TGGAGAACGATAATGGCGAAGCTGTGATCAAGAAGAGTTTCCTGGAGATA 1428

Qy    1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
            |||   |  | || ||    |||||| |  | ||     || | |||||||| | |   
Db    1429 GAGAAGAAGGTTATTACACATTGGAAGGATGTAGTGATGAACGGCAAGAAGACTAATAAA 1488

Qy    1501 ATCTGCTAA 1509
            | |||| |
Db    1489 GTTTGCTGA 1497


RESULT 11
US-10-167-547C-3
; Sequence 3, Application US/10167547C
; Publication No. US20030170653A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  APPLICANT:  Damude, Howard G.
;  TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
;  TITLE OF INVENTION:  Butyrolactone and its Intermediates
;  FILE REFERENCE: CL1804 US NA
;  CURRENT APPLICATION NUMBER: US/10/167,547C
;  CURRENT FILING DATE:  2003-03-17
;  PRIOR APPLICATION NUMBER: 60/297198
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: Microsoft Office 07
; SEQ ID NO 3
;   LENGTH: 1530
;   TYPE: DNA
;   ORGANISM: Alstroemeria
US-10-167-547C-3

  Query Match             49.8%;  Score 751;  DB 12;  Length 1530;
  Best Local Similarity   72.4%;  Pred. No. 9.6e-238;
  Matches  973;  Conservative    0;  Mismatches  370;  Indels    0;  Gaps    0;

Qy       1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
           ||||  ||||||  || ||| || || ||  ||  |   || || | ||| |  || || 
Db       1 ATGGCTCTCTCCAGCGTCGTCTCCGACTCCAACAACCAAGTGCAGTGCACCTATGCCTCT 60

Qy      61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
           || ||||| ||         |||  |||||| |||||||||| || ||||| || ||||| 
Db      61 CGCTACGTTCGCGACGAGGCTCCAGGGTTCAGGATGCCGGAGAAGTCGATACCAAAGGAG 120

Qy     121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
           |||||||    ||| |||||||||||||||||||| |||||||| ||  || ||||||| 
Db     121 GCGGCGTTCACGATGATCAACGACGAGCTGATGCTGGACGGGAACCCCAGGCTGAACTTG 180

Qy     181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
           || || || |||||||| ||||||||||| ||||| |||   || || ||||||| |||||
Db     181 GCTTCGTTCGTGACGACGTGGATGGAGCCGGAGTGCGATCGTCTGATGATGTCCACCATC 240

Qy     241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
           ||||||||||| |    ||||||||| ||||| ||||| |  ||  | ||||| || || 
Db     241 AACAAGAACTACGCCCTCATGGACGATTACCCGGTCACTATTGACATACAGAATCGCTGC 300

Qy     301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
           ||||| ||||| ||  | || || ||||| ||  | |  |||| ||| ||  | || |||
Db     301 GTGAATATGATAGCCAACCTCTTTAATGCGCCAATTGGGGAGGGGGAAACAACAGTAGGA 360

Qy     361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              |  || || |||||||| || ||||| ||| | || || ||||| |||||| | || 
```

```
        361 TGTGCTACGGTGGGATCATCAGAAGCCATGATGCTTGCAGGGTTGGCATTCAAGAGAAAT 420

        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
            |||||||||||  | || || || || || ||    || || |||||||| ||||||||
        421 TGGCAGAACAAAAGAAAGGCAGAGGGGAAGCCATATGACAAGCCCAACATGGTCACCGGT 480

        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
             | ||||||||| || || ||| |||||||||||| ||| ||||||||| ||||| | ||  ||
        481 TCAAATGTTCAGGTTTGCTGGGTGAAATTCGCTAAGTATTTTGAAGTTGAAATGAAAAAA 540

        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
            ||||| ||||| || ||||| ||||||||||||||||  |  | ||||| || ||||| |||
        541 GTGAATTTGAGGGAGGGATATTATGTGATGGACCCAGAGAAGGCTGTGGAAATGGTGGAT 600

        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
            ||||| ||||||||||||||| | |||  | || || || |||| |||||| |||||||||
        601 GAGAATACCATTTGTGTTGCTGCCATCTTGGGCTCGACCCTTACTGGAGAGTTCGAAGAT 660

        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
            || |||||  | ||||| ||| | || || ||||||||  |||| || |||||||||
        661 GTCAAACTGCTAAACGACCTCCTCGTAGAGAAGAACAAGAAAACTGGTTGGGATACACCC 720

        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
            || || || ||||| || | ||| |||||||||||| || ||| | ||||| ||| |||||
        721 ATTCATGTTGATGCTGCTATTGGTGGATTCATTGCTCCATTTATCTATCCAGAACTGGAA 780

        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
            ||||| ||  |||| ||  |||||||||||||||||| ||||| ||||| || || |||
        781 TGGGATTTCCGACTACCTCTGGTGAAGAGTATCAATGTCAGTGGACACAAATACGGCCTT 840

        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
            || ||  | ||  ||||||||||  ||||||| |||||  | ||| | ||||| ||||||
        841 GTCTATCCCGGCGTTGGTTGGGTCGTCTGGAGGAACAAGAATGATCTTCCTGAAGAACTC 900

y       901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
            || ||||||||||| |||||||||   ||| |||||||| |||| ||||| ||||| |||
b       901 ATTTTCCATATCAACTATCTTGGGATTGATCAACCCACTTTTACCCTCAACTTCTCGAAA 960

y       961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
            |||||||  ||  | |||| ||||||||| ||| | || ||  | ||    ||||| |||
b       961 GGTTCAAACCAGATAATTGGTCAATACTATCAATTAATTCGCCTTGGTTTTGAGGGGTAC 1020

y      1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
            |  ||| | ||||| || ||||  ||||||   |   | || ||||||   || ||| ||
b      1021 AAGAATATAATGGAAAACTGCACGGAGAATGCAAGAATTCTTAGGGAACATCTCGAGGAG 1080

y      1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
            |  |     ||| | ||| |||| |||||    || | ||| || ||| |  |  |  |
b      1081 ATGGGCGTTTTCGAGATCATCTCCAAGGATATTGGGGCGCCTCTCGTCACAATTGCTCTC 1140

y      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
            || || ||||||   || | ||  ||  | ||  |||| |   || | |||| |||||||
b      1141 AAGGACAGCAGCAAACATAGTGTCTTTAAGATAGCCGATACAATTAGAAGGTTTGGATGG 1200

y      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
            | | | || || ||||||||||||     | |   | |||||  | || ||||||||||||
b      1201 ACAATTCCTGCATACACAATGCCAAAGGACGTTGAGCACATAGCCGTCCTTCGTGTGGTT 1260

y      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
            ||||| || || |||    | |  ||||| ||| | || |  |  || ||  |||| |||
Db     1261 ATCAGGGAGGACTTCAGCCGGAGCCTCGCCGAGCGCCTAGCTAATGACATGAAGAAGGTG 1320

Qy     1321 ATGCGTGAGCTCGATGAGCTTCC 1343
             ||  ||||||| ||    | |||
Db     1321 TTGGTTGAGCTGGACATACATCC 1343

RESULT 12
US-10-167-547C-5
; Sequence 5, Application US/10167547C
; Publication No. US20030170653A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  APPLICANT:  Damude, Howard G.
;  TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
;  TITLE OF INVENTION:  Butyrolactone and its Intermediates
;  FILE REFERENCE: CL1804 US NA
;  CURRENT APPLICATION NUMBER: US/10/167,547C
;  CURRENT FILING DATE:  2003-03-17
;  PRIOR APPLICATION NUMBER: 60/297198
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: Microsoft Office 07
; SEQ ID NO 5
;   LENGTH: 1590
;   TYPE: DNA
;   ORGANISM: Alstroemeria
US-10-167-547C-5

  Query Match             49.8%;  Score 751;  DB 12;  Length 1590;
  Best Local Similarity   72.4%;  Pred. No. 9.8e-238;
  Matches  973;  Conservative    0;  Mismatches  370;  Indels    0;  Gaps    0;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||  ||||||  || ||| || || ||  ||  |   || || | ||| |  || ||
Db          1 ATGGCTCTCTCCAGCGTCGTCTCCGACTCCAACAACCAAGTGCAGTGCACCTATGCCTCT 60

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || ||||| ||         |||  ||||||| ||||||||| || ||||| || ||||| 
Db         61 CGCTACGTTCGCGACGAGGCTCCAGGGTTCAGGATGCCGGAGAAGTCGATACCAAAGGAG 120

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              |||||||    ||| ||||||||||||||||||| |||||||| ||  || |||||||
Db        121 GCGGCGTTCACGATGATCAACGACGAGCTGATGCTGGACGGGAACCCCAGGCTGAACTTG 180

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              || || || ||||||| ||||||||| ||||| |||   || || |||||| |||||
Db        181 GCTTCGTTCGTGACGACGTGGATGGAGCCGGAGTGCGATCGTCTGATGATGTCCACCATC 240

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||| |    ||||||||| ||||| ||||| |  ||  | ||||| || ||
Db        241 AACAAGAACTACGCCCTCATGGACGATTACCCGGTCACTATTGACATACAGAATCGCTGC 300

Qy        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              ||||| ||||| ||  | || || ||||| ||  | |  |||| ||| ||   | || |||
Db        301 GTGAATATGATAGCCAACCTCTTTAATGCGCCAATTGGGGAGGGGGAAACAACAGTAGGA 360

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
                 |  || || |||||||| || ||||| ||| | || || ||||| ||||| |||| | ||
Db        361 TGTGCTACGGTGGGATCATCAGAAGCCATGATGCTTGCAGGGTTGGCATTCAAGAGAAAT 420

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              |||||||||||  | || || || || || ||    || || |||||||| ||||||||
Db        421 TGGCAGAACAAAAGAAAGGCAGAGGGGAAGCCATATGACAAGCCCAACATGGTCACCGGT 480

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
               | ||||||||| || || ||| |||||||||||| ||| ||||| ||||| | ||  ||
Db        481 TCAAATGTTCAGGTTTGCTGGGTGAAATTCGCTAAGTATTTTGAAGTTGAAATGAAAAAA 540

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              ||||| ||||| || ||||| ||||||||||||||||  |  | ||||| || ||||| |||
Db        541 GTGAATTTGAGGGAGGGATATTATGTGATGGACCCAGAGAAGGCTGTGGAAATGGTGGAT 600

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              ||||| ||||||||||||||| | |||  | || || || |||| |||||| |||||||||
Db        601 GAGAATACCATTTGTGTTGCTGCCATCTTGGGCTCGACCCTTACTGGAGAGTTCGAAGAT 660

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              || |||||  | ||||| ||| | || || ||||||||  |||| || |||||||||
Db        661 GTCAAACTGCTAAACGACCTCCTCGTAGAGAAGAACAAGAAAACTGGTTGGGATACACCC 720

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || || || ||||| || | ||| |||||||||||| || ||| | ||||| ||| |||||
Db        721 ATTCATGTTGATGCTGCTATTGGTGGATTCATTGCTCCATTTATCTATCCAGAACTGGAA 780

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
```

```
              ||||| ||  |||| ||  ||||||||||||||||||| ||||| ||||| || || |||
          781 TGGGATTTCCGACTACCTCTGGTGAAGAGTATCAATGTCAGTGGACACAAATACGGCCTT 840

          841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || ||  | ||  ||||||||||  ||||||| |||||  | ||| | ||||| ||||||
          841 GTCTATCCCGGCGTTGGTTGGGTCGTCTGGAGGAACAAGAATGATCTTCCTGAAGAACTC 900

          901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              || ||||||||||| ||||||||   ||| |||||||| ||||| ||||| ||||| |||
          901 ATTTTCCATATCAACTATCTTGGGATTGATCAACCCACTTTTACCCTCAACTTCTCGAAA 960

          961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              |||||||  ||  | |||| ||||||||| ||| | || ||  | ||    ||||| |||
          961 GGTTCAAACCAGATAATTGGTCAATACTATCAATTAATTCGCCTTGGTTTTGAGGGGTAC 1020

,        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              |  ||| | ||||| || ||||  ||||||   |   | || ||||||   || ||| ||
ɔ        1021 AAGAATATAATGGAAAACTGCACGGAGAATGCAAGAATTCTTAGGGAACATCTCGAGGAG 1080

y        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              |  |     ||| | ||| |||| |||||    || | ||| || ||| |  |  |  |
ɔ        1081 ATGGGCGTTTTCGAGATCATCTCCAAGGATATTGGGGCGCCTCTCGTCACAATTGCTCTC 1140

y        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              || || ||||||   || | ||  ||  | ||  |||| |   || | |||| |||||||
b        1141 AAGGACAGCAGCAAACATAGTGTCTTTAAGATAGCCGATACAATTAGAAGGTTTGGATGG 1200

y        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              | | | || || |||||||||||     | |   | |||||  | || ||||||||||||
b        1201 ACAATTCCTGCATACACAATGCCAAAGGACGTTGAGCACATAGCCGTCCTTCGTGTGGTT 1260

y        1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              ||||| || || |||    | |  ||||| ||| | || |  |  || ||  |||| |||
b        1261 ATCAGGGAGGACTTCAGCCGGAGCCTCGCCGAGCGCCTAGCTAATGACATGAAGAAGGTG 1320

y        1321 ATGCGTGAGCTCGATGAGCTTCC 1343
               ||  |||||| ||    | |||
b        1321 TTGGTTGAGCTGGACATACATCC 1343

ESULT 13
S-10-006-852-17
 Sequence 17, Application US/10006852
 Publication No. US20030046732A1
 GENERAL INFORMATION:
  APPLICANT: Kinnersely, Alan M.
  APPLICANT:  Turano, Frank J.
  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
, FILE REFERENCE: 7224-65
, CURRENT APPLICATION NUMBER: US/10/006,852
, CURRENT FILING DATE:  2002-07-01
, PRIOR APPLICATION NUMBER: US 60/246,367
, PRIOR FILING DATE: 2000-11-07
, NUMBER OF SEQ ID NOS: 24
, SOFTWARE: PatentIn version 3.1
, SEQ ID NO 17
;   LENGTH: 1783
;   TYPE: DNA
;   ORGANISM: Lycopersicon esculentum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (6)..(1511)
;   OTHER INFORMATION:
US-10-006-852-17

  Query Match             48.8%;  Score 736.6;  DB 14;  Length 1783;
  Best Local Similarity   71.8%;  Pred. No. 6.4e-233;
  Matches  964;  Conservative    0;  Mismatches  379;  Indels    0;  Gaps    0;

Qy         20 TATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCAC 79
              ||   || || || |    | | ||||  ||||| |||||| | || || |      |
Db         28 TAAGAGATTCAGAAGAGAGCTTGCACTGTACATTTGCATCAAGATATGTACAGGAACCTT 87
```

```
Qy         80 TTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCA 139
              | |||| |||||| |||||  |||| || || || || ||||| || ||||||||  | |
Db         88 TACCTAAGTTCAAAATGCCTAAAAAATCCATGCCGAAAGAAGCAGCTTATCAGATTGTAA 147

Qy        140 ACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACAT 199
              |||||||||| ||| | || || || ||  ||||||| ||||| |||||||| |  ||||
Db        148 ACGACGAGCTTATGTTGGATGGTAACCCCAGGTTGAATTTAGCTTCCTTTGTTAGCACAT 207

Qy        200 GGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACA 259
              |||||||||| ||||| ||||| ||||||||||| ||||| || || |||||||| ||||
Db        208 GGATGGAGCCCGAGTGCGATAAGCTCATCATGTCATCCATTAATAAAAACTATGTCGACA 267

Qy        260 TGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATC 319
              |||| ||||| || |||||||| |||||||| ||  ||||||| |||||| | |||||||
Db        268 TGGATGAGTATCCTGTCACCACTGAACTTCAAAATAGATGTGTTAACATGTTAGCACATC 327

Qy        320 TATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCAT 379
              | ||| |||| ||| | |  || |  ||||| || || ||||| || || || || ||||
Db        328 TTTTCCATGCCCCGGTTGGTGATGATGAGACTGCAGTTGGAGTTGGTACAGTGGGTTCAT 387

Qy        380 CGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAG 439
              | ||||| |||||| | || ||  | || ||||| || ||||||||    ||  | ||||
Db        388 CAGAGGCAATAATGCTTGCTGGCCTTGCTTTCAAACGCAAATGGCAATCGAAAAGAAAAG 447

Qy        440 CTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTT 499
              | |||||||||||  ||||||| || || || ||||| ||||| ||||| || || || |
Db        448 CAGAAGGCAAACCTTTCGATAAGCCTAATATAGTCACTGGAGCTAATGTGCAGGTCTGCT 507

Qy        500 GGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGAT 559
              |||| ||||| || ||||| |||||||||||  | ||||| |||||| | |  |||||||
Db        508 GGGAAAAATTTGCAAGGTATTTTGAGGTTGAGTTGAAGGAGGTGAAACTAAAAGAAGGAT 567

Qy        560 ACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTG 619
              ||||||| |||||||||    |||| || || || || |||||||| || || |||||||
Db        568 ACTATGTAATGGACCCTGCCAAAGCAGTAGAGATAGTGGATGAGAATACAATATGTGTTG 627

Qy        620 CGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATC 679
              | |  || |||||||| ||||| | ||| || || || ||||| || ||| | ||||| |
Db        628 CTGCAATCCTTGGTTCTACTCTGACTGGGGAGTTTGAGGATGTGAAGCTCCTAAACGAGC 687

Qy        680 TCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAA 739
              || |     |||||||||| |||||||||||||| ||||| || || || ||||| || |
Db        688 TCCTTACAAAAAAGAACAAGGAAACCGGATGGGAGACACCGATTCATGTCGATGCTGCGA 747

Qy        740 GTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCT 799
              |||||||||| ||||| || ||  | |  || ||  | |||||||| ||  |  | ||
Db        748 GTGGAGGATTTATTGCTCCTTTCCTCTGGCCAGATCTTGAATGGGATTTCCGTTTGCCTC 807

Qy        800 TGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTT 859
              | ||||| ||||| ||||| || |||||||||||||| ||||| || || ||  | ||||
Db        808 TTGTGAAAAGTATAAATGTCAGCGGTCACAAGTATGGCCTTGTATATGCTGGTGTCGGTT 867

Qy        860 GGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATC 919
              ||||||| ||| | | ||| || || ||||| || |||||| |||| ||||| || || |
Db        868 GGGTGATATGGCGGAGCAAGGAAGACTTGCCCGATGAACTCGTCTTTCATATAAACTACC 927

Qy        920 TTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTG 979
              ||||  |||| || || || ||||||||||| ||||| ||||||||   |||| | ||||
Db        928 TTGGGTCTGATCAGCCTACTTTTACTCTCAACTTCTCTAAAGGTTCCTATCAAATAATTG 987

Qy        980 CTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATT 1039
              | || || || ||  | ||  || | |||   |||||||| |  || || ||| ||||||
Db        988 CACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGTTATAAGAACGTCATGAAGAATT 1047

Qy       1040 GCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCG 1099
              ||  |   ||    |  || ||||  || ||| |    || |  |   ||||| | || |
Db       1048 GCTTATCAAACGCAAAAGTACTAACAGAGGGAATCACAAAAATGGGGCGGTTCGATATTG 1107

Qy       1100 TCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACA 1159
              |||| ||||| | ||| || ||  |||| || || ||  | |  || ||||||    | |
Db       1108 TCTCTAAGGATGTGGGTGTTCCTGTTGTAGCATTTTCTCTCAGGGACAGCAGCAAATATA 1167

Qy       1160 CTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAA 1219
```

```
          1168 CGGTATTTGAAGTATCTGAGCATCTCAGAAGATTTGGATGGATCGTCCCTGCATACACAA 1227

          1220 TGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGA 1279

          1228 TGCCACCGGATGCTGAACACATTGCTGTACTGCGGGTTGTCATTAGAGAGGATTTCAGCC 1287

          1280 GAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGC 1339

          1288 ACAGCCTAGCTGAGAGACTTGTTTCTGACATTGAGAAAATTCTGTCAGAGTTGGACACAC 1347

          1340 TTCCTTCGAGAGTGATTCACAAA 1362

          1348 AGCCTCCTCGTTTGCCCACCAAA 1370


 SULT 14
 -09-727-758-1
 Sequence 1, Application US/09727758
 Publication No. US20030101487A1
 GENERAL INFORMATION:
  APPLICANT: KISAKA, HIROAKI
  APPLICANT:  KIDA, TAKAO
  TITLE OF INVENTION: A method for producing transgenic plants having improved amino ac
  TITLE OF INVENTION:  composition
  FILE REFERENCE: 200397US0
  CURRENT APPLICATION NUMBER: US/09/727,758
  CURRENT FILING DATE:  2000-12-04
  PRIOR APPLICATION NUMBER: JP 11-352552
  PRIOR FILING DATE: 1999-12-13
  NUMBER OF SEQ ID NOS: 5
  SOFTWARE: PatentIn version 3.1
 SEQ ID NO 1
   LENGTH: 1180
   TYPE: DNA
   ORGANISM: Lycopersicon esculentum
S-09-727-758-1

 Query Match            47.1%;  Score 711.4;  DB 11;  Length 1180;
 Best Local Similarity  75.7%;  Pred. No. 1.1e-224;
 Matches  894;  Conservative    0;  Mismatches  286;  Indels    1;  Gaps    1;

)y        112 CCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGG 171

)b          1 CCAAAGGAAGCAGCATATCAGATTCTAAATGATGAACTTATGTTAGATGGAAATCCAAGG 60

)y        172 TTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATG 231

Db         61 TTGAATTTGGCATCTTTTGTGACAACATGGATGGAACCAGAATGTGACAAATTGATGATG 120

Qy        232 TCCTCCATCAACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAG 291

Db        121 GATTCCATTAACAAAAATTATGTTGACATGGATGAATATCCTGTCACCACTGAGCTTCAG 180

Qy        292 AACCGATGTGTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACC 351

Db        181 AATCGGTGTGTAAACATGATAGCGCATTTATTTAATGCACCACTTGAAGATGGAGAAACT 240

Qy        352 GCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTC 411

Db        241 GCAGTTGGAGTTGGAACAGTTGGTTCTTCAGAAGCCATTATGCTTGCTGGATTGGCCTTT 300

Qy        412 AAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATT 471

Db        301 AAGAGAAAATGGCAAAACAAAATGAAAGCCCAAGGAAAGCCCTATGATAAGCCCAACATT 360

Qy        472 GTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAA 531

Db        361 GTTACTGGTGCTAATGTCCAGGTGTGTTGGGAAAAATTTGCAAGGTATTTTGAAGTTGAG 420

Qy        532 CTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGAT 591

Db        421 CTAAAAGAAGTGAAGTTGACTGATGGATACTATGTGATGAGCCCTGAGAAAGCTGTGGAA 480
```

```
Qy        592 ATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAA 651

Db        481 ATGGTTGATGAGAACACAATTTGTGTAGCTGCTATTTTGGGTTCAACTCTGAATGGGGAG 540

Qy        652 TTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGG 711

Db        541 TTTGAGGAAGTGAAGAAATTGAATGACCTCCT-ATTGAAAAGAACAAGGAAACAGGGTGG 599

Qy        712 GATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCG 771

Db        600 GACACTCCAATTCATGTGGATGCAGCTAGTGGTGGATTTATTGCACCATTTATATATCCA 659

Qy        772 GAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAG 831

Db        660 GAACTTGAATGGGACTTTAGATTGCCATTAGTTAAAAGTATAAATGTCAGCGGTCACAAG 719

Qy        832 TATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCT 891

Db        720 TATGGCCTTGTATATGCTGGTGTCGGTTGGGTCATATGGCGGAGCAAGGAAGACTTGCCC 779

Qy        892 GAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAAT 951

Db        780 GATGAACTTATTTTTCATATTAATTACCTTGGGTCTGATCAACCTACTTTCACCCTCAAT 839

Qy        952 TTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCAC 1011

Db        840 TTCTCTAAAGGATCAAGTCAAGTAATTGCTCAATATTATCAACTAATTCGTTTTGGTTAT 899

Qy       1012 GAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGA 1071

Db        900 GAGGGTTATAAGAACGTCATGAAGAATTGCTTATCAAACGCAAGGGTACTAACAGAGGGA 959

Qy       1072 CTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCT 1131

Db        960 ATCACAAAAATGGGAAGATTCGAGATTGTCTCCAAGGATGTTGGTGTTCCCTTGGTTGCA 1019

Qy       1132 TTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGG 1191

Db       1020 TTTTCTCTTAAAGACAATAGCAAACATGATGAGTTTGAAGTATCTGAAACTTTAAGGAGA 1079

Qy       1192 TATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTT 1251

Db       1080 TTTGGATGGATTGTTCCAGCATACACTATGCCACCAGATGCTCAACACATCACAGTTCTT 1139

Qy       1252 CGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCTGA 1292

Db       1140 AGAGTTGTTATTAGAGAAGATTTCTCCCGTACACTCGCCGA 1180


RESULT 15
US-09-887-576-782
; Sequence 782, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
;  APPLICANT: Budworth, P.
;  APPLICANT:  Brown, D.
;  APPLICANT:  Chang, H.
;  APPLICANT:  Zhu, T.
;  APPLICANT:  Han, B.
;  APPLICANT:  Wang, X.
;  APPLICANT:  Cooper, Bret
;  TITLE OF INVENTION: Promoters for regulation of plant expression
;  FILE REFERENCE: 1360.001US1
;  CURRENT APPLICATION NUMBER: US/09/887,576
;  CURRENT FILING DATE:  2001-06-25
;  PRIOR APPLICATION NUMBER: US 60/213,848
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/214,087
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/258,692
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 875
```

```
 SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 782
  LENGTH: 1455
  TYPE: DNA
  ORGANISM: Oryza sativa
-09-887-576-782

Query Match             47.0%;  Score 709.6;  DB 10;  Length 1455;
Best Local Similarity   71.8%;  Pred. No. 5.1e-224;
Matches  928;  Conservative    0;  Mismatches  364;  Indels    0;  Gaps    0;

/            45 CTCCACATTCGCATCACGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAA 104
                ||| || ||||| || || ||||| || ||  | || || ||||||| ||||||||| ||
)            42 CTCGACGTTCGCGTCGCGGTACGTGCGCACGGCGCTGCCGAGGTTCAGGATGCCGGAGAA 101

y           105 CTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAA 164
                 ||||| || ||||| ||||||||| |||||||||||||||||||||||||| ||||| ||
)           102 GTCGATCCCCAAGGACGCGGCGTACCAGATCATCAACGACGAGCTGATGCTCGACGGCAA 161

y           165 TCCACGGTTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACT 224
                 || ||  ||||| | || ||||| || || || ||||||||||| ||||| || || ||
b           162 CCCGCGCCTGAACCTGGCGTCCTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCT 221

y           225 CATCATGTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGA 284
                ||| ||| || |||||||||||||||| || |||||||| ||||||||||||||||||||
b           222 CATGATGGCCGCCATCAACAAGAACTACGTCGACATGGATGAGTACCCCGTCACCACCGA 281

y           285 ACTTCAGAACCGATGTGTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGC 344
                 || |||||||| || ||||||||||| || ||||| ||||| || ||| | |  || |
b           282 GCTCCAGAACCGGTGCGTGAACATGATCGCGCATCTGTTCAACGCGCCGATCGGGGACGA 341

y           345 GGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTT 404
                 ||||| || ||||| || || || || || || ||||||||||| ||| |||| ||  |
)b          342 CGAGACGGCGGTCGGGGTGGGCACGGTGGGGTCGTCGGAGGCCATCATGCTGGCGGGGCT 401

)y          405 GGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACC 464
                ||| |||||| | || |||||||||| |   || || || || || |||  ||| || ||
)b          402 GGCGTTCAAGAGGAAGTGGCAGAACAGGATGAAGGCCGAGGGGAAGCCCCACGACAAGCC 461

Qy          465 CAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGA 524
                |||||| || || || ||||| || || ||||| |||||||| |||||  | ||||| ||
Db          462 CAACATCGTGACGGGGGCCAACGTGCAGGTGTGCTGGGAGAAGTTCGCGCGCTACTTCGA 521

Qy          525 GGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGC 584
                ||| || || ||||| |||||  |||   |||| ||||| |||||| ||||  |  | ||
Db          522 GGTGGAGCTCAAGGAGGTGAAGCTGACCCAAGGGTACTACGTGATGAACCCGGAGAAGGC 581

Qy          585 TGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAA 644
                 || || ||||| || ||||||||||| || || || | ||| || || ||||| || ||
Db          582 CGTGGAGATGGTCGACGAGAACACCATCTGCGTCGCCGCCATCCTCGGCTCCACCCTCAA 641

Qy          645 TGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAAC 704
                 || || ||||| || || ||  |  | ||||| ||| |   ||  ||||||   || ||
Db          642 CGGCGAGTTCGAGGACGTCAAGATGCTCAACGACCTCCTCACCGCCAAGAACGCCGAGAC 701

Qy          705 CGGATGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTT 764
                 || ||| | || || ||||| ||||| ||||| || || || ||||| || |||||  |
Db          702 AGGGTGGAACACGCCGATCCATGTGGACGCGGCGAGCGGCGGGTTCATCGCGCCGTTCAT 761

Qy          765 GTATCCGGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGG 824
                 || |||||  |||| ||||||||  | || || ||||||||||| ||||| || || ||
Db          762 CTACCCGGAGCTGGAGTGGGACTTCCGGCTGCCGCTGGTGAAGAGCATCAACGTCAGCGG 821

Qy          825 TCACAAGTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGA 884
                 |||||||| || || || ||||| ||  | || ||||| |||||| | ||||| |||||
Db          822 CCACAAGTACGGGCTCGTCTACGCCGGCGTCGGGTGGGTCATCTGGCGCAACAAGGAGGA 881

Qy          885 TTTGCCTGAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTAC 944
                  | || || || ||||||||||| ||||| || || || || ||||| || ||||| ||
Db          882 CCTCCCCGATGAGCTCATCTTCCACATCAACTACCTCGGCGCCGACCAGCCAACCTTCAC 941

Qy          945 TCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATT 1004
                 ||||| ||||||||||| || |  ||  | ||||| || || ||||| || || ||  |
Db          942 GCTCAACTTCTCCAAAGGATCGAACCAGATAATTGCGCAGTATTACCAGCTCATTCGTCT 1001

Qy         1005 GGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAG 1064
                 ||   |||||| ||||   |  | ||| |||| ||| | || ||   ||  || ||  |
Db         1002 CGGATTCGAGGGGTACAAGGACATCATGCAGAACTGCCGGGACAACGCGACGGTGCTCCG 1061

Qy         1065 GGAAGGACTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACT 1124
                ||| ||  | |||||||| |     ||| || | || || ||||||   || ||||| ||
Db         1062 GGAGGGGATCGAGAAGACGGGCCACTTCGACGTGGTGTCCAAGGACTCCGGCGTGCCGCT 1121

Qy         1125 TGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCT 1184
                 || || |||||| | || ||        |  |||| | ||||||  |  |||| |  ||
Db         1122 GGTGGCCTTCTCCCTCAAGGACTCGTCGCGGTACACGGTGTTCGAGGTGGCCGAGAGCCT 1181

Qy         1185 TCGCAGGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCAC 1244
                 ||| | |  || ||||| |||||||| ||||| |||||  |  | ||  | ||| || |
Db         1182 CCGCCGCTTCGGCTGGATCGTGCCGGCGTACACCATGCCCGCCGACGCTGAGCACGTCGC 1241

Qy         1245 TGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGAT 1304
                 ||  | || || || ||| | || || |||    |    ||||| ||| | ||  | |
Db         1242 CGTGATGCGCGTCGTCATCCGCGAGGACTTCAGCCGCGGCCTCGCCGAGCGCCTCATCAC 1301

Qy         1305 CGATATAGAGAAAGTGATGCGTGAGCTCGATG 1336
                |||  |    ||   | ||   ||  | || |
Db         1302 CGACCTCACCAAGACGGTGGCCGATATGGACG 1333

Search completed: October 22, 2003, 15:04:15
Job time : 439 secs
```

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2003  Compugen Ltd.

 nucleic - nucleic search, using sw model

 n on:          October 22, 2003, 10:35:12 ; Search time 117 Seconds
                                          (without alignments)
                                          5692.711 Million cell updates/sec

 tle:           US-10-006-852-1
 rfect score:   1509
 quence:        1 atggtgctctcccacgccgt..........agacgagtggtatctgctaa 1509

 oring table:   IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

 arched:        569978 seqs, 220691566 residues

 tal number of hits satisfying chosen parameters:      1139956

 nimum DB seq length: 0
 aximum DB seq length: 2000000000

 st-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

 atabase :       Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                 %
esult          Query
 No.    Score  Match Length DB  ID                        Description
-------------------------------------------------------------------------
    1   736.6   48.8    1783  2  US-08-522-421-8           Sequence 8, Appli
c   2   311.2   20.6 4403765  3  US-09-103-840A-2           Sequence 2, Appli
c   3   311.2   20.6 4411529  3  US-09-103-840A-1           Sequence 1, Appli
    4   270.6   17.9    5565  3  US-09-068-195-21          Sequence 21, Appl
c   5     133    8.8   36941  4  US-08-311-731A-130        Sequence 130, App
c   6      64    4.2     359  4  US-09-056-556-192         Sequence 192, App
c   7      64    4.2     359  4  US-09-072-596-187         Sequence 187, App
c   8    41.6    2.8    7218  1  US-08-232-463-14          Sequence 14, Appl
c   9    39.2    2.6 1230025  4  US-09-198-452A-1           Sequence 1, Appli
c  10    35.8    2.4   16442  3  US-08-781-891-208         Sequence 208, App
c  11    35.8    2.4   16442  4  US-09-618-166-208         Sequence 208, App
c  12    35.6    2.4   15016  4  US-09-601-198-60          Sequence 60, Appl
c  13    33.6    2.2     515  4  US-09-026-408-11          Sequence 11, Appl
   14    33.4    2.2    2004  1  US-08-471-033-18          Sequence 18, Appl
   15    33.4    2.2    2004  2  US-08-471-044-18          Sequence 18, Appl
   16    33.4    2.2    2004  2  US-08-463-483A-18         Sequence 18, Appl
   17    33.4    2.2    2004  2  US-08-471-046A-18         Sequence 18, Appl
   18    33.4    2.2    2004  2  US-08-470-566B-18         Sequence 18, Appl
   19    33.4    2.2    2004  2  US-08-469-334-18          Sequence 18, Appl
   20    33.4    2.2    2004  3  US-09-300-529-18          Sequence 18, Appl
   21    33.4    2.2    2576  1  US-08-471-033-35          Sequence 35, Appl
   22    33.4    2.2    2576  2  US-08-471-044-35          Sequence 35, Appl
   23    33.4    2.2    2576  2  US-08-463-483A-35         Sequence 35, Appl
   24    33.4    2.2    2576  2  US-08-471-046A-35         Sequence 35, Appl
   25    33.4    2.2    2576  2  US-08-470-566B-35         Sequence 35, Appl
   26    33.4    2.2    2576  2  US-08-469-334-35          Sequence 35, Appl
   27    33.4    2.2    2576  3  US-09-300-529-35          Sequence 35, Appl
```

```
   28    33.4    2.2    2655  1  US-08-471-033-17          Sequence 17, Appl
   29    33.4    2.2    2655  1  US-08-471-033-26          Sequence 26, Appl
   30    33.4    2.2    2655  2  US-08-471-044-17          Sequence 17, Appl
   31    33.4    2.2    2655  2  US-08-471-044-26          Sequence 26, Appl
   32    33.4    2.2    2655  2  US-08-463-483A-17         Sequence 17, Appl
   33    33.4    2.2    2655  2  US-08-463-483A-26         Sequence 26, Appl
   34    33.4    2.2    2655  2  US-08-471-046A-17         Sequence 17, Appl
   35    33.4    2.2    2655  2  US-08-471-046A-26         Sequence 26, Appl
   36    33.4    2.2    2655  2  US-08-470-566B-17         Sequence 17, Appl
   37    33.4    2.2    2655  2  US-08-470-566B-26         Sequence 26, Appl
   38    33.4    2.2    2655  2  US-08-469-334-17          Sequence 17, Appl
   39    33.4    2.2    2655  2  US-08-469-334-26          Sequence 26, Appl
   40    33.4    2.2    2655  3  US-09-300-529-17          Sequence 17, Appl
   41    33.4    2.2    2655  3  US-09-300-529-26          Sequence 26, Appl
   42    33.4    2.2    4031  1  US-08-471-033-49          Sequence 49, Appl
   43    33.4    2.2    4031  2  US-08-471-044-49          Sequence 49, Appl
   44    33.4    2.2    4031  2  US-08-463-483A-49         Sequence 49, Appl
   45    33.4    2.2    4031  2  US-08-471-046A-49         Sequence 49, Appl

                                  ALIGNMENTS


RESULT 1
US-08-522-421-8
; Sequence 8, Application US/08522421
; Patent No. 5908973
;  GENERAL INFORMATION:
;    APPLICANT:  Abu-Bakar, Umi Kalsom
;    APPLICANT:  Barton, Sarah Louise
;    APPLICANT:  Gallego-Veigas, Pedro Pablo
;    APPLICANT:  Gray, Julie Elizabeth
;    APPLICANT:  Grierson, Donald
;    APPLICANT:  Lowe, Alexandra Louise
;    APPLICANT:  Picton, Steve
;    APPLICANT:  Whotton, Lee Colin
;    TITLE OF INVENTION:  DNA, DNA CONSTRUCTS, CELLS AND PLANTS
;    TITLE OF INVENTION:  DERIVED THEREFROM
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pillsbury Madison & Sutro, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3918
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Microsoft Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/522,421
;      FILING DATE:  11-JAN-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9305868.3
;      FILING DATE:  22-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9305869.1
;      FILING DATE:  22-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9305859.2
;      FILING DATE:  22-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9305865.9
;      FILING DATE:  22-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9305866.7
;      FILING DATE:  22-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9305867.5
;      FILING DATE:  22-MAR-1993
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9305860.0
;      FILING DATE:  22-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9305862.6
;      FILING DATE:  22-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9314351.9
;      FILING DATE:  12-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9320988.0
;      FILING DATE:  12-OCT-1993
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1783 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;      ORGANISM:  ERTD1
US-08-522-421-8

  Query Match             48.8%;  Score 736.6;  DB 2;  Length 1783;
  Best Local Similarity   71.8%;  Pred. No. 4.2e-232;
  Matches  964;  Conservative    0;  Mismatches  379;  Indels    0;  Gaps    0;

Qy          20 TATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCAC 79
               ||   || || || |    | | ||||  ||||| |||||| | || || |      |
Db          28 TAAGAGATTCAGAAGAGAGCTTGCACTGTACATTTGCATCAAGATATGTACAGGAACCTT 87

Qy          80 TTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCA 139
               | |||| |||||| |||||  |||| || || || || ||||| || ||||||||  | |
Db          88 TACCTAAGTTCAAAATGCCTAAAAAATCCATGCCGAAAGAAGCAGCTTATCAGATTGTAA 147

Qy         140 ACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACAT 199
               |||||||||| ||| | || || || ||  ||||||| ||||| |||||||| |  ||||
Db         148 ACGACGAGCTTATGTTGGATGGTAACCCCAGGTTGAATTTAGCTTCCTTTGTTAGCACAT 207

Qy         200 GGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACA 259
               |||||||||| ||||| ||||| ||||||||||| ||||| || || |||||||| ||||
Db         208 GGATGGAGCCCGAGTGCGATAAGCTCATCATGTCATCCATTAATAAAAACTATGTCGACA 267

Qy         260 TGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATC 319
               |||| ||||| || |||||||| |||||||| ||  ||||||| |||||| | |||||||
Db         268 TGGATGAGTATCCTGTCACCACTGAACTTCAAAATAGATGTGTTAACATGTTAGCACATC 327

Qy         320 TATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCAT 379
               | ||| |||| ||| | |  || |  ||||| || || ||||| || || || || ||||
Db         328 TTTTCCATGCCCCGGTTGGTGATGATGAGACTGCAGTTGGAGTTGGTACAGTGGGTTCAT 387

Qy         380 CGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAG 439
               | ||||| |||||| | || ||  | || ||||| || ||||||||    ||  | ||||
Db         388 CAGAGGCAATAATGCTTGCTGGCCTTGCTTTCAAACGCAAATGGCAATCGAAAAGAAAAG 447

Qy         440 CTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTT 499
               | |||||||||||  ||||||| || || || ||||| ||||| ||||| || || || |
Db         448 CAGAAGGCAAACCTTTCGATAAGCCTAATATAGTCACTGGAGCTAATGTGCAGGTCTGCT 507

Qy         500 GGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGAT 559
               |||| ||||| || |||||||||||||||||  | ||||| |||||| | |  |||||||
Db         508 GGGAAAAATTTGCAAGGTATTTTGAGGTTGAGTTGAAGGAGGTGAAACTAAAAGAAGGAT 567

Qy         560 ACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTG 619
               ||||||| |||||||||    |||| || || || || |||||||| || || |||||||
Db         568 ACTATGTAATGGACCCTGCCAAAGCAGTAGAGATAGTGGATGAGAATACAATATGTGTTG 627

Qy         620 CGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATC 679
               | |  || |||||||| ||||| | ||| || || || ||||| || ||| | ||||| |
Db         628 CTGCAATCCTTGGTTCTACTCTGACTGGGGAGTTTGAGGATGTGAAGCTCCTAAACGAGC 687

Qy         680 TCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAA 739
               || |     |||||||||| |||||||||||||| ||||| || || || ||||| || |
Db         688 TCCTTACAAAAAAGAACAAGGAAACCGGATGGGAGACACCGATTCATGTCGATGCTGCGA 747

Qy         740 GTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCT 799
               |||||||||| ||||| || ||  | |  || ||  | |||||||| ||  |  | ||
Db         748 GTGGAGGATTTATTGCTCCTTTCCTCTGGCCAGATCTTGAATGGGATTTCCGTTTGCCTC 807

Qy         800 TGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTT 859
               | ||||| ||||| ||||| || |||||||||||||| ||||| || || ||  | ||||
Db         808 TTGTGAAAAGTATAAATGTCAGCGGTCACAAGTATGGCCTTGTATATGCTGGTGTCGGTT 867

Qy         860 GGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATC 919
               ||||||| ||| | | ||| || || ||||| || |||||| |||| ||||| || || |
Db         868 GGGTGATATGGCGGAGCAAGGAAGACTTGCCCGATGAACTCGTCTTTCATATAAACTACC 927

Qy         920 TTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTG 979
               ||||  |||| || || || ||||||||||| ||||| ||||||||   |||| | ||||
Db         928 TTGGGTCTGATCAGCCTACTTTTACTCTCAACTTCTCTAAAGGTTCCTATCAAATAATTG 987

Qy         980 CTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATT 1039
               | || || || ||  | ||  || | |||   |||||||| |  || || ||| ||||||
Db         988 CACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGTTATAAGAACGTCATGAAGAATT 1047

Qy        1040 GCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCG 1099
               ||  |   ||    |  || ||||  || ||| |    || |  |   ||||| | || |
Db        1048 GCTTATCAAACGCAAAAGTACTAACAGAGGGAATCACAAAAATGGGGCGGTTCGATATTG 1107

Qy        1100 TCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACA 1159
               |||| ||||| | ||| || ||  |||| || || ||  | |  || ||||||    | |
Db        1108 TCTCTAAGGATGTGGGTGTTCCTGTTGTAGCATTTTCTCTCAGGGACAGCAGCAAATATA 1167

Qy        1160 CTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAA 1219
               | |  || ||| | || ||    ||  | || | ||||||||| || || || |||||||
Db        1168 CGGTATTTGAAGTATCTGAGCATCTCAGAAGATTTGGATGGATCGTCCCTGCATACACAA 1227

Qy        1220 TGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGA 1279
               |||| ||  ||||  |||||||  |||| || || || || || ||||| ||||||
Db        1228 TGCCACCGGATGCTGAACACATTGCTGTACTGCGGGTTGTCATTAGAGAGGATTTCAGCC 1287

Qy        1280 GAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGC 1339
                 |  || ||||||||||||||    || || |||||| |  ||   ||| | ||    |
Db        1288 ACAGCCTAGCTGAGAGACTTGTTTCTGACATTGAGAAAATTCTGTCAGAGTTGGACACAC 1347

Qy        1340 TTCCTTCGAGAGTGATTCACAAA 1362
                 ||| |  |  ||     ||||
Db        1348 AGCCTCCTCGTTTGCCCACCAAA 1370

RESULT 2
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
 Query Match             20.6%;  Score 311.2;  DB 3;  Length 4403765;
 Best Local Similarity   54.1%;  Pred. No. 1e-89;
 Matches  687;  Conservative    0;  Mismatches  568;  Indels   15;  Gaps    2;

y           80 TTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCA 139
               | ||   | |   |||||| ||  | |||||   |   || ||||| ||||   |||||
b      3837572 TGCCGGCGCTGCGGATGCCCGATGAGTCGATGGATCCCGAGGCGGCCTATCGCTTCATCC 3837513

y          140 ACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACAT 199
               |||||||||||||||| ||||| |   | ||| ||||  | ||| |||| ||||| || |
b      3837512 ACGACGAGCTGATGCTCGACGGTAGCTCGCGGCTGAATCTGGCCACCTTCGTGACCACCT 3837453

y          200 GGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACA 259
               ||||||| || |||   || || || ||        |  || ||||||||    | || |
b      3837452 GGATGGACCCGGAGGCCGAAAAGCTGATGGCCGAGACGTTCGACAAGAACATGATCGATA 3837393

y          260 TGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATC 319
                |||||| ||||| |  ||| | |   |  ||   || ||||||  |||| | ||  | |
b      3837392 AGGACGAATACCCGGCGACCGCGGCCATCGAGGCGCGCTGTGTGTCCATGGTCGCCGACC 3837333

y          320 TATTCAATGC---------ACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCG 370
               | ||  | ||          |      | ||  |    | ||||  ||| ||    |||
b      3837332 TGTTTCACGCCGAGGGTCTGCGCGACCACGACCCCACCAGCGCCACCGGGGTGTCCACCA 3837273

y          371 TTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACA 430
               | || ||    |||||  | ||| |||  ||  ||||| | ||  |    |||| | |
Db     3837272 TCGGCTCCAGCGAGGCGGTGATGCTGGGTGGGCTGGCCCTGAAATGGCGTTGGCGGCAAC 3837213

Qy         431 AGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTC 490
                |       ||         | |      |  |||||  | || |  ||  | || || |
Db     3837212 GGGTGGGGTCTTGGAAGGGGCGC------ATGCCCAATCTGGTGATGGGTTCGAACGTCC 3837159

Qy         491 AAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGA 550
               | |||   ||||||||| |||    | ||||| || || ||||      |  ||    ||
Db     3837158 AGGTGGTGTGGGAGAAGTTCTGCCGCTACTTCGACGTCGAACCCCGTTACCTGCCGATGG 3837099

Qy         551 GTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCA 610
                    ||   ||| || ||   ||| | ||| |   | |     || || ||||||||||
Db     3837098 AGCGGGGCCGCTACGTCATCACCCCCGAGCAGGTGCTCGCCGCCGTCGACGAGAACACCA 3837039

Qy         611 TTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCT 670
               |  | || | ||  ||  | ||  ||||   ||  || ||| |||||    |
Db     3837038 TCGGCGTGGTGGCGATCTTGGGCACCACCTATACCGGTGAACTCGAACCCATCGCCGAGA 3836979

Qy         671 TGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGG 730
               |   ||      ||| | ||  |      |     ||   |||    ||  | ||||| |
Db     3836978 TCTGCGCCGCGCTGGACAAACTGGCGGCTGGCGGGGGTGTGGACGTCCCGGTACACGTCG 3836919

Qy         731 ATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTA 790
               | ||||| ||||| || ||  | |  ||||||||| |||||||  ||| |||||| |||
Db     3836918 ACGCGGCCAGTGGGGGCTTTGTGGTGCCGTTTTTGCATCCGGACCTGGTATGGGATTTTC 3836859

Qy         791 GACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAG 850
               | || |||    |||  |   ||||| || || || |||||||||||| ||    ||  | |
Db     3836858 GGCTGCCCCGCGTGGTGTCGATCAACGTCAGCGGCCACAAGTATGGGCTGACCTATCCCG 3836799

Qy         851 GGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATA 910
               |  | || |  ||  | ||| |       ||| |  |||| ||||| ||  | ||||
Db     3836798 GCGTCGGGTTTGTCGTGTGGCGCGGGCCCGAGCACCTGCCGGAGGATCTGGTTTTCCGGG 3836739

Qy         911 TCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTC 970
               |||| || || || |  |||   || ||||| || || || ||||||         |  |
Db     3836738 TCAACTACCTCGGCGGCGACATGCCGACCTTCACCCTGAACTTCTCCCGTCCCGGTAACC 3836679

Qy         971 AAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGA 1030
               | ||  | |  || |||||| |  |  | ||  |||| | ||| || || |  || ||||
Db     3836678 AGGTGGTGGGCCAGTACTACAACTTCCTGCGGCTGGGGCGCGACGGCTATACCAAGGTGA 3836619

Qy        1031 TGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGT 1090
               || ||          | |           ||  | ||    ||     ||   ||    |
Db     3836618 TGCAGGCGCTGTCGCACACCGCCCGGTGGCTGGGTGACCAGCTGCGCGAGGTGGATCATT 3836559

Qy        1091 TCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCA 1150
                | |  |  ||||  | |    ||   | ||  | |||   |||    |    |    |
Db     3836558 GCGAGGTGATCTCGGATGGTTCGGCGATCCCGGTGGTCAGCTTCCGGCTCGCCGGCGACC 3836499

Qy        1151 GCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGG 1210
               || |  |||| ||||||||  ||||| |   ||| || |  |  || |||   ||||| |
Db     3836498 GCGGGTACACGGAGTTCGACGTCTCCCATGAGCTGCGGACCTTCGGGTGGCAGGTGCCCG 3836439

Qy        1211 CCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAG 1270
               ||||||| |||||    || ||     || |  | || || ||  | ||  |  | ||||
Db     3836438 CCTACACCATGCCGGACAACGCCACCGACGTGGCGGTGCTGCGGATCGTGGTTCGGGAAG 3836379

Qy        1271 ATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGC 1330
                  ||||       || ||   |   ||      |||    |  |  | | ||  ||  |
Db     3836378 GACTCTCCGCCGACCTGGCGCGGGCCCTGCACGACGACGCGGTCACCGCGTTGGCTGCCC 3836319

Qy        1331 TCGATGAGCT 1340
               | ||  || |
Db     3836318 TGGACAAGGT 3836309


RESULT 3
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             20.6%;  Score 311.2;  DB 3;  Length 4411529;
  Best Local Similarity   54.1%;  Pred. No. 1e-89;
  Matches  687;  Conservative    0;  Mismatches  568;  Indels   15;  Gaps    2;

Qy          80 TTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCA 139
               | ||   | |   |||||| ||  | |||||   |   || ||||| ||||   |||||
Db     3851674 TGCCGGCGCTGCGGATGCCCGATGAGTCGATGGATCCCGAGGCGGCCTATCGCTTCATCC 3851615

Qy         140 ACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACAT 199
               |||||||||||||||| ||||| |   | ||| ||||  | ||| |||| ||||| || |
Db     3851614 ACGACGAGCTGATGCTCGACGGTAGCTCGCGGCTGAATCTGGCCACCTTCGTGACCACCT 3851555

Qy         200 GGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACA 259
               ||||||| || |||   || || || ||        |  || ||||||||    | || |
Db     3851554 GGATGGACCCGGAGGCCGAAAAGCTGATGGCCGAGACGTTCGACAAGAACATGATCGATA 3851495

Qy         260 TGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATC 319
                |||||| ||||| |  ||| | |   |  ||   || ||||||  |||| | ||  | |
Db     3851494 AGGACGAATACCCGGCGACCGCGGCCATCGAGGCGCGCTGTGTGTCCATGGTCGCCGACC 3851435

Qy         320 TATTCAATGC---------ACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCG 370
               | ||  | ||          |      | ||  |    | ||||  ||| ||    |||
Db     3851434 TGTTTCACGCCGAGGGTCTGCGCGACCACGACCCCACCAGCGCCACCGGGGTGTCCACCA 3851375

Qy         371 TTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACA 430
               | || ||    |||||  | ||| |||  ||  ||||| | ||  |    |||| | |
Db     3851374 TCGGCTCCAGCGAGGCGGTGATGCTGGGTGGGCTGGCCCTGAAATGGCGTTGGCGGCAAC 3851315

Qy         431 AGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTC 490
                |       ||         | |      |  |||||  | || |  ||  | || || |
```

```
      3851314 GGGTGGGGTCTTGGAAGGGGCGC------ATGCCCAATCTGGTGATGGGTTCGAACGTCC 3851261

          491 AAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGA 550
              | |||   |||||||| |||    | ||||| || || ||||      |  ||    ||
      3851260 AGGTGGTGTGGGAGAAGTTCTGCCGCTACTTCGACGTCGAACCCCGTTACCTGCCGATGG 3851201

          551 GTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCA 610
                   ||   ||| || ||   |||  |  || |   | |     || || ||||||||||
      3851200 AGCGGGGCCGCTACGTCATCACCCCCGAGCAGGTGCTCGCCGCCGTCGACGAGAACACCA 3851141

          611 TTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCT 670
              |  | || | ||  ||  | ||  ||||   ||  || ||| |||||    |
      3851140 TCGGCGTGGTGGCGATCTTGGGCACCACCTATACCGGTGAACTCGAACCCATCGCCGAGA 3851081

          671 TGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGG 730
              |   ||      ||| | ||  |      |     ||   |||    ||  | ||||| |
      3851080 TCTGCGCCGCGCTGGACAAACTGGCGGCTGGCGGGGGTGTGGACGTCCCGGTACACGTCG 3851021

/         731 ATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTA 790
              | ||||| ||||| || ||  | |  ||||||||| ||||||| ||| |||||| |||
ɔ     3851020 ACGCGGCCAGTGGGGGCTTTGTGGTGCCGTTTTTGCATCCGGACCTGGTATGGGATTTTC 3850961

y         791 GACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAG 850
              | || |||   |||  |   ||||| || || || ||||||||||| ||    ||  | |
b     3850960 GGCTGCCCCGCGTGGTGTCGATCAACGTCAGCGGCCACAAGTATGGGCTGACCTATCCCG 3850901

y         851 GGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATA 910
              |  | || |  ||  | ||| |       ||| |  |||| ||||| ||  | ||||
b     3850900 GCGTCGGGTTTGTCGTGTGGCGCGGGCCCGAGCACCTGCCGGAGGATCTGGTTTTCCGGG 3850841

y         911 TCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTC 970
              |||| || || || |  |||   || ||||| || || || ||||||         |  |
b     3850840 TCAACTACCTCGGCGGCGACATGCCGACCTTCACCCTGAACTTCTCCCGTCCCGGTAACC 3850781

y         971 AAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGA 1030
              | ||  | |  || |||||| |  |  | ||  |||| | ||| || || |  || ||||
b     3850780 AGGTGGTGGGCCAGTACTACAACTTCCTGCGGCTGGGGCGCGACGGCTATACCAAGGTGA 3850721

y        1031 TGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGT 1090
              || ||          | |           ||  | ||    ||     ||   ||    |
)b    3850720 TGCAGGCGCTGTCGCACACCGCCCGGTGGCTGGGTGACCAGCTGCGCGAGGTGGATCATT 3850661

Qy       1091 TCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCA 1150
               | |  |  ||||  | |    ||   | ||  | |||   |||    |    |    |
Db    3850660 GCGAGGTGATCTCGGATGGTTCGGCGATCCCGGTGGTCAGCTTCCGGCTCGCCGGCGACC 3850601

Qy       1151 GCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGG 1210
              || |  |||| ||||||||  ||||| |   ||| || |  |  || |||   ||||| |
Db    3850600 GCGGGTACACGGAGTTCGACGTCTCCCATGAGCTGCGGACCTTCGGGTGGCAGGTGCCCG 3850541

Qy       1211 CCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAG 1270
              ||||||| |||||    || ||     || |  | || || ||  | ||  |  | ||||
Db    3850540 CCTACACCATGCCGGACAACGCCACCGACGTGGCGGTGCTGCGGATCGTGGTTCGGGAAG 3850481

Qy       1271 ATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGC 1330
                 ||||       || ||   |   ||      |||    |  |  | | ||  ||  |
Db    3850480 GACTCTCCGCCGACCTGGCGCGGGCCCTGCACGACGACGCGGTCACCGCGTTGGCTGCCC 3850421

Qy       1331 TCGATGAGCT 1340
              | ||  || |
Db    3850420 TGGACAAGGT 3850411

RESULT 4
US-09-068-195-21
; Sequence 21, Application US/09068195B
; Patent No. 6140078
; GENERAL INFORMATION:
;  APPLICANT: Sanders, Jan W.
;  APPLICANT: Ledeboer, Adrianus M.
;  APPLICANT: Venema, Gerard
;  APPLICANT: Kok, Jan
;  TITLE OF INVENTION: Salt-Inducible Promoter Derivable from a Lactic Acid
;  TITLE OF INVENTION: Bacterium, and Its Use in a Lacctic Acid Bacterium for
;  TITLE OF INVENTION: Production of Desired Protein
;  FILE REFERENCE: Sanders-60113/0252227
;  CURRENT APPLICATION NUMBER: US/09/068,195B
;  CURRENT FILING DATE: 1998-07-29
;  EARLIER APPLICATION NUMBER: PCT/EP97/04755
;  EARLIER FILING DATE: 1997-08-20
;  EARLIER APPLICATION NUMBER: EP 97200744/7
;  EARLIER FILING DATE: 1997-03-13
;  EARLIER APPLICATION NUMBER: EP 96202444/4
;  EARLIER FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 5565
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: genomic DNA of
;   OTHER INFORMATION: NS3 locus containing gadRCB
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (3)..(500)
;   OTHER INFORMATION: C-terminus of rnhB
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1095)..(1922)
;   OTHER INFORMATION: rggL = gadR
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (2069)..(3577)
;   OTHER INFORMATION: orfX = gadC
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (3600)..(4997)
;   OTHER INFORMATION: gadB
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: Complement((5078)..(5563))
;   OTHER INFORMATION: C-terminus of ORF in opposite direction
US-09-068-195-21

  Query Match             17.9%;  Score 270.6;  DB 3;  Length 5565;
  Best Local Similarity   52.4%;  Pred. No. 3e-78;
  Matches  678;  Conservative    0;  Mismatches  599;  Indels   18;  Gaps    3;

Qy         80 TTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCA 139
              | ||||  |  ||  |  |  || | || |||        || ||| |||||| |  |
Db       3679 TACCTAAATATAAATTAGCTCAACAATCAATTGAGCCTCGAGTGGCCTATCAGTTAGTTC 3738

Qy        140 ACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACAT 199
              | || ||  || |   ||| |||||  | || || || || ||| | ||       || |
Db       3739 AAGATGAAATGCTAGATGAAGGGAACGCTCGTTTAAATTTGGCCACATTCTGTCAAACTT 3798

Qy        200 GGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACA 259
                ||||| |||||    |  || || || |      || |  | || ||     ||||||
Db       3799 ATATGGAACCTGAAGCAGTCAAGCTGATGAGTCAGACCTTGGAAAAAAATGCGATTGACA 3858

Qy        260 TGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATC 319
                   || || ||    || || ||| || | ||||| || || |||||||| ||  | |
Db       3859 AATCAGAATATCCAAGAACAACTGAAATTGAAAACCGTTGCGTCAACATGATCGCTGACC 3918

Qy        320 TATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCAT 379
              | |  |||||      | ||      |  |        ||       ||  |||| || |
Db       3919 TTTGGAATGCGAGTGAAAAA------GGAAAAATTTATGGGACTTCGACAATTGGTTCTT 3972

Qy        380 CGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAG 439
              | || ||    ||| | |  ||  ||||  | |||  |   ||||  ||        ||
Db       3973 CAGAAGCTTGTATGCTTGGGGGAATGGCTATGAAGTTTTCTTGGCGTAAGCGAGCAGAAA 4032

Qy        440 CTGAAGGCAAACCCGTCGAT------AAACCCAACATTGTCACCGGAGCCAATGTTCAAG 493
                  ||||  |    |  ||      || || ||| | ||||      |   |  |||||
Db       4033 AATTAGGCCTAGATATTAATGCGAAAAAGCCAAACTTAGTCATTTCCTCTGGTTATCAAG 4092
```

```
       494 TGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTG 553
           | || ||||| ||||||  |   || |  ||  ||||| | |  ||||||  | ||  |
      4093 TTTGCTGGGAAAAATTCTGTGTTTATTGGGATATTGAAATGAGAGAAGTGCCAATGGATA 4152

       554 AAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTT 613
            || | |       ||  |     || |||   | |||   |||||||| | || |||
      4153 GAGAACATATGTCAATCAATTTGGAAAAAGTGATGGATTATGTTGATGAATATACGATTG 4212

       614 GTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGA 673
           | || |  |  ||| | ||    |||  || |||    |  || ||| ||||   |||
      4213 GAGTAGTTGGAATTATGGGGATTACTTATACTGGTCGTTATGATGATATCAAAGCTTTGG 4272

'      674 ACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATG 733
           |  || |  | |  ||| | || ||| | || |  |   |       || ||||| ||||
)     4273 ATAATTTGATTGAAGAATATAATAAACAGACAGACTACAAAGTTTATATTCACGTAGATG 4332

'      734 CGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGAC 793
           | ||    |||||| |   ||| || ||| |  | || ||| | || |||||| ||  |
)     4333 CTGCTTCAGGAGGACTTTATGCTCCTTTTGTTGAGCCAGAACTTGAGTGGGATTTCCGTT 4392

y      794 TTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGA 853
           |       || |     ||||||      || || || ||||||  | || ||  | ||
)     4393 TGAAAAATGTCATTTCAATCAATACTTCAGGACATAAATATGGTTTAGTATATCCTGGTG 4452

y      854 TTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCA 913
           | ||||||||  | ||| |  ||||| |  |||| ||||| ||  | || ||  |  | |
D     4453 TAGGTTGGGTCTTGTGGCGTGACAAAAAATATTTACCTGAAGAGTTAATTTTTAAAGTAA 4512

y      914 ATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAG 973
            ||||||||| |  ||   ||| ||  |  |  | ||||| ||  |  || |   ||||
b     4513 GTTATCTTGGAGGAGAATTACCAACAATGGCGATTAATTTTTCTCACAGTGCTTCTCAAT 4572

y      974 TCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGG 1033
           | || | |||||||||  |  || | || |  ||    || || || | |  | |    |
b     4573 TAATCGGTCAATACTATAATTTTGTACGTTATGGATTTGATGGATATAAAGCTATTCATG 4632

y     1034 AGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCA 1093
           |||         |       ||     ||   |||| | |||| || ||| | || |||
)b    4633 AGAGAACGCATAAAGTAGCCATGTATTTAGCAGAAGAAATTGAAAAAACAGGAATGTTTG 4692

2y    1094 ACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGA------TA 1147
           | ||  |     | |  |    | | ||| |||||   | |   || |||||       |
)b    4693 AGATTATGAACGATGGGGCACAATTACCAATTGTCTGCTACAAATTAAAAGAAAATTCAA 4752

2y    1148 GCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGC 1207
            | |  ||   | |   |  ||  |  | ||     |    | | | ||||||  |||||
)b    4753 ACCGTGGTTGGAATCTTTATGATTTGGCAGATCGTTTATTAATGAAGGGATGGCAAGTGC 4812

Qy    1208 CGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAG 1267
           | || ||  || | |||  ||||    || |    |   |||  ||| | || ||  |||
Db    4813 CTGCTTATCCACTTCCTAAAAATTTGGAAAATGAAATCATTCAACGTTTAGTAATTCGAG 4872

Qy    1268 AAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTG 1327
            |||||||  ||  |   | ||    |    |||    ||||  |  |||  ||   ||
Db    4873 CAGATTTCGGGATGAATATGGCATTTAACTATGTTCAAGATATGCAAGAAGCAATTGATG 4932

Qy    1328 AGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAA 1362
             ||  || ||  || |     | |   |||  ||
Db    4933 CACTAAATAAGGCTCATATTCTATTTCATCAGGAA 4967
```

RESULT 5
US-08-311-731A-130/c

```
; Sequence 130, Application US/08311731A
; Patent No. 6583266
;  GENERAL INFORMATION:
;    APPLICANT:  SMITH, DOUGLAS
;    APPLICANT:  MAO, JEN-I
;    TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES
;    TITLE OF INVENTION:  RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
;    TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;    NUMBER OF SEQUENCES:  411
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WOLF, GREENFIELD & SACKS, P.C.
;      STREET:  600 ATLANTIC AVENUE
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,731A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GATES, EDWARD R.
;      REGISTRATION NUMBER:  31,616
;      REFERENCE/DOCKET NUMBER:  C0044/7125
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/720-3500
;      TELEFAX:  617/720-2441
;  INFORMATION FOR SEQ ID NO:  130:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36941 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  MYCOBACTERIUM LEPRAE
US-08-311-731A-130

  Query Match             8.8%;  Score 133;  DB 4;  Length 36941;
  Best Local Similarity   53.3%;  Pred. No. 2.3e-32;
  Matches  511;  Conservative    0;  Mismatches  425;  Indels   23;  Gaps   10;

Qy          93 GATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTGAT 152
               | |||| |||   |||||   |   || ||||||| |   | ||  ||||| |||| ||
Db       10476 GCTGCCCGAAGCATCGATGGATCCTGAGGCGGCGTACCGCTTTATTCACGACTAGCTAAT 10417

Qy         153 GCTTGACGGGAATCCACGGTTGAACTTAGCCT--CCTTTGTGACGACATGGATGGAGCCT 210
               | | ||||| | ||   || | ||| ||||    | || || || || || |||| ||
Db       10416 GGTAGACGGCAGTTCTCGGCTCAACCTAGCGCTACTTTCGTCACCACCTGGGTGGATCCC 10357

Qy         211 GAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTAC 270
               |||    | ||  | ||        |  || |||| |  | | | |||| |||||||||
Db       10356 GAGGTAAAGAAGTTGATGGCCGAGACGTTCGACAAAACATGTTTAGACAAGGACGAGTAT 10297

Qy         271 CCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATCTATTCAATGCA 330
               || |   ||||  |    |||| |   || | | |  || ||        || || ||
Db       10296 CCGG---CGGCCGCGCCATCGAACGGCGCTGCGTATAGCATGGCGACCGGTCTCTATTCA 10240

Qy         331 CCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATA 390
               |   |||  ||  | |  ||  |   | |   |  | | || ||  ||      |
Db       10239 CGCCTAG-GGACTCAGCTACGACAACCCGTCCAATACCTGTGGGGTGTTCCATTATCGGC 10181

Qy         391 ATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAA 450
                    ||  || | |   | |  |  |   ||||   | ||  ||   | ||   |  |
Db       10180 TCCAGGGATGGGTGGTGATGCTGGGTTGCCTGGCCATAAAATTGCGCTGGTGTGCGGTAG 10121

Qy         451 CCCGT--------CGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGG 502
               |  ||            |  |||||  | |  |  || ||    || || |||  |||||
Db       10120 CTAGTTAGCAGCGTAGCACCCCCAATCTGGGGATGGGCGCAGGGGTCCAGGTGGTTTGGG 10061

Qy         503 AGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACT 562
               |||| |||  | | ||||| || || || |      |  ||    ||   ||||||| ||
Db       10060 AGAAGTTCTGTCGTTACTTCGATGTCGAGC--CGCTATCTGTCGATGGAGGAAGGATGCT 10003

Qy         563 ATGTGATGGACCCTCAACAAGCTGTTGATATG--GTTGATGAGAACACCATTTGTGTTGC 620
```

```
           | || ||   | | | |  || |  |        || || ||| |||| | |||   |
     10002 ACGTCATCACCTCCCGAGCAGGTAATCGAGGCCCGTCGACGAGGACACTACTTGGTGTCG 9943

       621 GGACATTCTTGGT--TCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGAT 678
            | |  |||| ||   ||||| | |  || ||| |||||   ||        |   ||
      9942 TGGCCATCTTTGTCTGCCACTTTCACCGGCGAACTCGAACCCGTCGCGGAGCTCTGCGTC 9883

       679 CTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCA 738
               ||| | |   |      || |  || | ||| |   || |   ||| | ||| ||
)     9882 GCGCTGGACAAGCTGGCAGCCGACAATGGCTTGGACATTGCAGTGTCCGTCGGTGCCGCC 9823

/      739 AGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCC 798
           || || || |||      || |  ||| | ||||| ||| | ||||| |   | || |||
)     9822 AGCGG-GGGTTCGGGATGCCTTGCTTGCACCCGGAGTTGAAGTGGGATTCCTGGCTGCCC 9764

/      799 TTGGTGAAGAGTATCAATGT-GAGTGGTCACAAGTATGGACT-TGTGTACGCAGGGATTG 856
              |||| |   ||||||||  |||||  | || ||||| ||     ||| | ||  | |
)     9763 CGAGTGATGTCGATCAATGTAAAGTGGCTATAAATATGGGCTCACCCTACCCCGGTGTGG 9704

y      857 GTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATT 916
           | |  |   | ||| |   ||| ||  ||||||| |||||  |  |||| | | |||| |
b     9703 GATTTGGCGTATGGCGCGGCAACGAATATTTGCCCGAGGACGTAGTCTTTCGTGTCAACT 9644

y      917 ATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCA 976
           | || || |  ||       || ||     | || || |||         |  |||||
b     9643 ACCTGGGCGGCGATATGTAGACTTTCGTATTGAACTTTTCCCGCCTGGACAACCAAGTGG 9584

y      977 TTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAG 1035
           ||  ||| |||||| |  || |  |  |||| |   | |||||| | || || ||| ||
b     9583 TTAGTCAGTACTACAACTTTCTGTGGCTGGGACGTAACGGTTACTGGAAAGTCATGCAG 9525

ESULT 6
S-09-056-556-192/c
 Sequence 192, Application US/09056556
 Patent No. 6350456
  GENERAL INFORMATION:
    APPLICANT:  Reed, Steven G.
    APPLICANT:  Skeiky, Yasir A.W.
    APPLICANT:  Dillon, Davin C.
    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND            TREATM
    NUMBER OF SEQUENCES:  241
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SEED and BERRY LLP
      STREET:  6300 Columbia Center, 701 Fifth Avenue
      CITY:  Seattle
      STATE:  Washington
      COUNTRY:  USA
      ZIP:  98104-7092
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/056,556
;     FILING DATE:  07-APR-1998
;     CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Maki, David J.
;     REGISTRATION NUMBER:  31,392
;     REFERENCE/DOCKET NUMBER:  210121.457
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (206) 622-4900
;     TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  192:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  359 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
US-09-056-556-192
```

```
  Query Match             4.2%;  Score 64;  DB 4;  Length 359;
  Best Local Similarity   58.3%;  Pred. No. 5.9e-11;
  Matches  112;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy        844 TACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATC 903
              ||  | ||  | || |  ||  | ||| |       ||| |  |||| ||||| ||  |
Db        353 TATCCCGGCGTCGGGTTTGTCGTGTGGCGCGGGCCCGAGCACCTGCCGGAGGATCTGGTT 294

Qy        904 TTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGT 963
              ||||   |||| || || || |  |||   || ||||| || || || ||||||
Db        293 TTCCGGGTCAACTACCTCGGCGGCGACATGCCGACCTTCACCCTGAACTTCTCCCGTCCC 234

Qy        964 TCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGA 1023
                 |  || ||  | |  || |||||| |  |  | ||  |||| | ||| || || |
Db        233 GGTAACCAGGTGGTGGGCCAGTACTACAACTTCCTGCGGCTGGGGCGCGACGGCTATACC 174

Qy       1024 AATGTGATGGAG 1035
              || |||||| ||
Db        173 AAGGTGATGCAG 162

RESULT 7
US-09-072-596-187/c
; Sequence 187, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  187:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-187

  Query Match             4.2%;  Score 64;  DB 4;  Length 359;
  Best Local Similarity   58.3%;  Pred. No. 5.9e-11;
  Matches  112;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy        844 TACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATC 903
```

```
         ||  | ||  | || |  ||  | ||| |       ||| |  |||| ||||| ||  |
     353 TATCCCGGCGTCGGGTTTGTCGTGTGGCGCGGGCCCGAGCACCTGCCGGAGGATCTGGTT 294

     904 TTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGT 963
         ||||   |||| || || || |  |||   || ||||| || || || ||||||
     293 TTCCGGGTCAACTACCTCGGCGGCGACATGCCGACCTTCACCCTGAACTTCTCCCGTCCC 234

     964 TCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGA 1023
            |  || ||  | |  || |||||| |  |  | ||  |||| | ||| || || |
     233 GGTAACCAGGTGGTGGGCCAGTACTACAACTTCCTGCGGCTGGGGCGCGACGGCTATACC 174

    1024 AATGTGATGGAG 1035
         || |||||| ||
     173 AAGGTGATGCAG 162


SULT 8
-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
 GENERAL INFORMATION:
   APPLICANT:  DORNER, F.
   APPLICANT:  SCHEIFLINGER, F.
   APPLICANT:  FALKNER, F. G.
   TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
   NUMBER OF SEQUENCES:  52
   CORRESPONDENCE ADDRESS:
     ADDRESSEE:  Foley & Lardner
     STREET:  1800 Diagonal Road, Suite 500
     CITY:  Alexandria
     STATE:  VA
     COUNTRY:  USA
     ZIP:  22313-0299
   COMPUTER READABLE FORM:
     MEDIUM TYPE:  Floppy disk
     COMPUTER:  IBM PC compatible
     OPERATING SYSTEM:  PC-DOS/MS-DOS
     SOFTWARE:  PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER:  US/08/232,463
     FILING DATE:
     CLASSIFICATION:  435
   PRIOR APPLICATION DATA:
     APPLICATION NUMBER:  US/07/935,313
     FILING DATE:
     APPLICATION NUMBER:  EP 91 114 300.6
     FILING DATE:  26-AUG-1991
   ATTORNEY/AGENT INFORMATION:
     NAME:  BENT, Stephen A.
     REGISTRATION NUMBER:  29,768
,     REFERENCE/DOCKET NUMBER:  30472/114 IMMU
,    TELECOMMUNICATION INFORMATION:
,      TELEPHONE:  (703)836-9300
,      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             2.8%;  Score 41.6;  DB 1;  Length 7218;
  Best Local Similarity   5.0%;  Pred. No. 0.01;
  Matches   20;  Conservative  208;  Mismatches  172;  Indels    0;  Gaps    0;

Qy         187 TTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAG 246
               ||||  ||:: : ::: ::::   ::: : :: :::   :  : :      :  :: :::
Db        1442 TTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1383
```

```
Qy         247 AACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAAC 306
               ::  : :  :: : ::: ::: :    :  :  :  :::    ::::  :: : : :::
Db        1382 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1323

Qy         307 ATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGA 366
               : ::  : : :   :   :: : :  :  :::::::: :::::  :  :  :::: ::::
Db        1322 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1263

Qy         367 ACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAG 426
               :  :  :::  :  :::::  : :: :  ::  ::   ::     ::: : ::: :: ::
Db        1262 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1203

Qy         427 AACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAAT 486
               :: ::: : ::::  ::::: :::   :  :: :::   :: :  :  :  ::::  ::
Db        1202 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1143

Qy         487 GTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAA 546
               :   ::: : :  ::::::::   :  ::: :    ::::  :::   ::::::: ::::
Db        1142 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1083

Qy         547 TTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTG 586
                 ::: :::::: :  ||   | |  ||||| ||  || |
Db        1082 RRRRRRRRRRRRRRRRATCGCAAGCTCCCTCGACCTGCAG 1043


RESULT 9
US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
;  APPLICANT: Griffais, R.
;  TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
;  TITLE OF INVENTION:  thereof and uses thereof, in particular for the diagnosis, preve
;  TITLE OF INVENTION:  and treatment of infection
;  FILE REFERENCE: 9710-003-999
;  CURRENT APPLICATION NUMBER: US/09/198,452A
;  CURRENT FILING DATE:  1998-11-24
;  NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
;   LENGTH: 1230025
;   TYPE: DNA
;   ORGANISM: Chlamydia pneumoniae
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(15000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (15001)..(30000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (30001)..(45000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (45001)..(60000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (60001)..(75000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (75001)..(90000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (90001)..(105000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (105001)..(120000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (120001)..(135000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (135001)..(150000)
;   OTHER INFORMATION: n=a or c or g or t
```

```
  NAME/KEY: misc_feature
  LOCATION: (150001)..(165000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (165001)..(180000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (180001)..(195000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (195001)..(210000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (210001)..(225000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (225001)..(240000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (240001)..(255000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (255001)..(270000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (270001)..(285000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (285001)..(300000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (300001)..(315000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (315001)..(330000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (330001)..(345000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (345001)..(360000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (360001)..(375000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (375001)..(390000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (390001)..(405000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
, LOCATION: (405001)..(420000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc_feature
, LOCATION: (420001)..(435000)
, OTHER INFORMATION: n=a or c or g or t
, NAME/KEY: misc_feature
; LOCATION: (435001)..(450000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (450001)..(465000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (465001)..(480000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (480001)..(495000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
; LOCATION: (495001)..(510000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc_feature
```

```
;   LOCATION: (510001)..(525000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (525001)..(540000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (540001)..(555000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (555001)..(570000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (570001)..(585000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (585001)..(600000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (600001)..(615000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (615001)..(630000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (630001)..(645000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (645001)..(660000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (660001)..(675000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (675001)..(690000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (690001)..(705000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (705001)..(720000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (720001)..(735000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (735001)..(750000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (750001)..(765000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (765001)..(780000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (780001)..(795000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (795001)..(810000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (810001)..(825000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (825001)..(840000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (840001)..(855000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (855001)..(870000)
;   OTHER INFORMATION: n=a or c or g or t
;   NAME/KEY: misc_feature
;   LOCATION: (870001)..(885000)
```

```
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (885001)..(900000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature
  LOCATION: (900001)..(915000)
  OTHER INFORMATION: n=a or c or g or t
  NAME/KEY: misc_feature

 Query Match             2.6%;  Score 39.2;  DB 4;  Length 1230025;
 Best Local Similarity   51.1%;  Pred. No. 1.8;
 Matches   92;  Conservative    0;  Mismatches   88;  Indels    0;  Gaps    0;

y         523 GAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAA 582
              ||  ||  |  || |      |||||  ||  |   |   ||| |  | | |    || |
b      784726 GAAATTTTAGATACGTTCAAAAAATTAGGTATACCTTTAGATGAGCAGAAGCGCTTACTA 784667

y         583 GCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTT 642
                ||| || |  |||| ||   |     ||| ||   |||  |  | ||||  ||||  ||
b      784666 AATGTCGAGAATGTTGCTGTAGATTTAGTTTTTGACTCGGTCTCTATTGGAACCACATTT 784607

y         643 AATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAA 702
              || | ||  || ||| | |    | |  |   |  |   ||| ||||  ||   |||||
b      784606 AAAGAAGCGTTGGAAAAGGCCGGAGTTATTTTCTGTTCTTTGGGCGAAGCGATTCAAGAA 784547


ESULT 10
S-08-781-891-208/c
 Sequence 208, Application US/08781891
 Patent No. 6090620
  GENERAL INFORMATION:
    APPLICANT:  Fu, Ying-Hui
    APPLICANT:  Yu, Chang-En
    APPLICANT:  Oshima, Junko
    APPLICANT:  Mulligan, John T.
    APPLICANT:  Schellenberg, Gerald D.
    TITLE OF INVENTION:  GENE AND GENE PRODUCTS RELATED TO
    TITLE OF INVENTION:  WERNER'S SYNDROME
    NUMBER OF SEQUENCES:  209
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SEED and BERRY LLP
      STREET:  6300 Columbia Center, 701 Fifth Avenue
      CITY:  Seattle
      STATE:  Washington
      COUNTRY:  USA
      ZIP:  98104-7092
    COMPUTER READABLE FORM:
,     MEDIUM TYPE:  Floppy disk
,     COMPUTER:  IBM PC compatible
,     OPERATING SYSTEM:  PC-DOS/MS-DOS
,     SOFTWARE:  PatentIn Release #1.0, Version #1.30
,   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/781,891
;     FILING DATE:  27-DEC-1996
;     CLASSIFICATION:  800
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  No. 6090620tenburg Ph.D., Carol
;     REGISTRATION NUMBER:  39,317
;     REFERENCE/DOCKET NUMBER:  240052.419
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (206) 622-4900
;     TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  208:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16442 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-781-891-208

  Query Match             2.4%;  Score 35.8;  DB 3;  Length 16442;
  Best Local Similarity   56.3%;  Pred. No. 1.4;
  Matches   67;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;
```

```
Qy        933 ACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTACCA 992
              ||||| |   |||||||||| |   | | |   ||  |||   ||| |   | | |||||
Db        207 ACCCAACCAAACTCTCAATTGCCATAGATGGAGAAAACAAAGTATTCCATGACAATACCA 148

Qy        993 ACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAATA 1051
              | ||  |  |||    | | | |  | |||   |   | ||| |||  |||| | || |
Db        147 AATTCACACATTATCTCTCCACGAATCCAGCCCTTCAAAGGATAATAACAGAAAAAAAA 89


RESULT 11
US-09-618-166-208/c
; Sequence 208, Application US/09618166
; Patent No. 6583112
;  GENERAL INFORMATION:
;       APPLICANT: Fu, Ying-Hui
;                  Yu, Chang-En
;                  Oshima, Junko
;                  Mulligan, John T.
;                  Schellenberg, Gerald D.
;       TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
;                           WERNER'S SYNDROME
;       NUMBER OF SEQUENCES: 209
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Seed Intellectual Property Law Group
;            STREET: 701 Fifth Avenue, Suite 6300
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98104-7092
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/618,166
;            FILING DATE: 17-Jul-2000
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McMasters, David D.
;            REGISTRATION NUMBER: 33,963
;            REFERENCE/DOCKET NUMBER: 240052.419C1
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 208:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 16442 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 208:
US-09-618-166-208

  Query Match             2.4%;  Score 35.8;  DB 4;  Length 16442;
  Best Local Similarity   56.3%;  Pred. No. 1.4;
  Matches   67;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy        933 ACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTACCA 992
              ||||| |   |||||||||| |   | | |   ||  |||   ||| |   | | |||||
Db        207 ACCCAACCAAACTCTCAATTGCCATAGATGGAGAAAACAAAGTATTCCATGACAATACCA 148

Qy        993 ACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAATA 1051
              | ||  |  |||    | | | |  | |||   |   | ||| |||  |||| | || |
Db        147 AATTCACACATTATCTCTCCACGAATCCAGCCCTTCAAAGGATAATAACAGAAAAAAAA 89


RESULT 12
US-09-601-198-60/c
; Sequence 60, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
```

```
 APPLICANT: Cassell, Gail H.
 APPLICANT:  Chen, Ellson Y.
 APPLICANT:  Glass, Jennifer S.
 APPLICANT:  Glass, John I.
 APPLICANT:  Heiner, Cheryl R.
 APPLICANT:  Lefkowitz, Elliot
 TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
 TITLE OF INVENTION:  UREALYTICUM
 FILE REFERENCE: UAB-13452/22
 CURRENT APPLICATION NUMBER: US/09/601,198
 CURRENT FILING DATE:  2000-12-08
 PRIOR APPLICATION NUMBER: 60/073,189
 PRIOR FILING DATE: 1998-01-30
 NUMBER OF SEQ ID NOS: 181
 SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 60
  LENGTH: 15016
  TYPE: DNA
  ORGANISM: Ureaplasma urealyticum
S-09-601-198-60

 Query Match             2.4%;  Score 35.6;  DB 4;  Length 15016;
 Best Local Similarity   47.7%;  Pred. No. 1.6;
 Matches  104;  Conservative    0;  Mismatches  114;  Indels    0;  Gaps    0;

y         461 AACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACT 520
              |||  ||  |||  | |   |    || | | || |     || ||     |     | |
b        9593 AACTAAAATTTGCTATCAATGATTTTGATGATGTTTTAAAAGAAAATCAAACCTTTAAGT 9534

y         521 TTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAAC 580
              || |  || |||  || |||   ||      | |   | |  |  | |  || | | ||
b        9533 TTAATATTCAACCAAATGAAAATAATGGTGCTAATTTAAATGAGATTAGCGAACATGAAG 9474

y         581 AAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTC 640
                   || || ||  |||||| | | |   | | ||||||  |  ||  |      |
b        9473 GTAAAGTAGAAATTATTGATGGGCAAAAATTCTTTGTTGCTAATCTTAATAACATTAAAG 9414

y         641 TTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGAT 678
              ||||   |||||   ||| |  |||| |||     |||
b        9413 TTAACAAAGAATATAAAGTTTATAAAATCTACTTTGAT 9376

.ESULT 13
JS-09-026-408-11/c
 Sequence 11, Application US/09026408
 Patent No. 6303338
  GENERAL INFORMATION:
    APPLICANT:  Ni et al.
    TITLE OF INVENTION:  PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
    TITLE OF INVENTION:  INHIBITOR
    NUMBER OF SEQUENCES:  15
    CORRESPONDENCE ADDRESS:
,     ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
,     STREET:  1100 NEW YORK AVENUE, SUITE 600
,     CITY:  WASHINGTON
,     STATE:  DC
,     COUNTRY:  USA
,     ZIP:  20005-3934
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/026,408
;     FILING DATE:  Herewith
;     CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/934,011
;     FILING DATE:  15-AUG-1997
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 60/024,056
;     FILING DATE:  16-AUG-1996
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0300002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-026-408-11

  Query Match             2.2%;  Score 33.6;  DB 4;  Length 515;
  Best Local Similarity   50.0%;  Pred. No. 0.78;
  Matches   81;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy        525 GGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGC 584
              ||||    | ||| |    ||  |   || ||  |||| | || |  ||  | | |  |
Db        506 GGTTATGTTCAAGNATCCAAACGTCTTTGGAGTCTACTTTTTGGTCTACTTTAATCTGGT 447

Qy        585 TGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAA 644
                |  | |     ||  |||| |   ||        | ||||| ||| | ||| | ||
Db        446 ATTNAAAAAAAAAGAAAAGAAAAGTCTTGAAACACTGCACATTTTTGCTACCAATTTTTT 387

Qy        645 TGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGT 686
                  ||   |  ||| | |||  | ||| |  |||| ||| |
Db        386 AATGGATGACTTAGAGGATAATATATTGTAAAATCTTTTGTT 345

RESULT 14
US-08-471-033-18
; Sequence 18, Application US/08471033
; Patent No. 5770696
;  GENERAL INFORMATION:
;    APPLICANT:  Warren, Gregory W
;    APPLICANT:  Koziel, Michael G
;    APPLICANT:  Mullins, Martha A
;    APPLICANT:  Nye, Gordon J
;    APPLICANT:  Carr, Brian
;    APPLICANT:  Desai, Nalini M
;    APPLICANT:  Kostichka, N. Kristy
;    APPLICANT:  Duck, Nicholas B
;    APPLICANT:  Estruch, Juan J
;    TITLE OF INVENTION:  No. 5770696el Pesticidal Proteins and Strains
;    NUMBER OF SEQUENCES:  50
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CIBA-GEIGY Corporation
;      STREET:  7 Skyline Drive
;      CITY:  Hawthorne
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/471,033
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/314,594
;      FILING DATE:  09-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/218,018
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
```

```
      FILING DATE:  25-MAR-1993
    ATTORNEY/AGENT INFORMATION:
      NAME:  Pace, Gary M.
      REGISTRATION NUMBER:  P-40,403
      REFERENCE/DOCKET NUMBER:  CGC 1695/CIP3/DIV7 - SQLv3
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  919-541-8582
      TELEFAX:  919-541-8689
  INFORMATION FOR SEQ ID NO:  18:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  2004 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    FEATURE:
      NAME/KEY:  misc_feature
      LOCATION:  1..2004
      OTHER INFORMATION:  /note= "Maize optimized DNA
      OTHER INFORMATION:  sequence for VIP1A(a) 80 kd protein from AB78"
3-08-471-033-18

 Query Match            2.2%;  Score 33.4;  DB 1;  Length 2004;
 Best Local Similarity   55.7%;  Pred. No. 2.2;
 Matches   64;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

y        1375 CAAGAGAAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATC 1434
              || ||||||    |    ||| ||||  |||  ||   |  |   | ||||||||| |
b        1357 CAGGAGAAGCTGAACAAGAACCGCGACTACTACATCAGCCTGTACATGAAGAGCGAGAAG 1416

y        1435 GACAAGCAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGA 1489
               |||  ||| | || |||| ||  | | |  |||  |    |  |  ||  ||||
b        1417 AACACCCAGTGCGAGATCACCATCGACGGCGAGATATACCCCATCACCACCAAGA 1471


ESULT 15
S-08-471-044-18
 Sequence 18, Application US/08471044
 Patent No. 5840868
  GENERAL INFORMATION:
    APPLICANT:  Warren, Gregory W
    APPLICANT:  Koziel, Michael G
    APPLICANT:  Mullins, Martha A
    APPLICANT:  Nye, Gordon J
    APPLICANT:  Carr, Brian
    APPLICANT:  Desai, Nalini M
    APPLICANT:  Kostichka, N. Kristy
    APPLICANT:  Duck, Nicholas B
    APPLICANT:  Estruch, Juan J
    TITLE OF INVENTION:  No. 5840868el Pesticidal Proteins and Strains
    NUMBER OF SEQUENCES:  50
    CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  CIBA-GEIGY Corporation
;     STREET:  7 Skyline Drive
;     CITY:  Hawthorne
;     STATE:  NY
;     COUNTRY:  USA
;     ZIP:  10532
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/471,044
;     FILING DATE:  06-JUN-1995
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/463,483
;     FILING DATE:  05-JUN-1995
;   PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 08/314,594
;      FILING DATE:  09-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/218,018
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/037,057
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Gary M.
;      REGISTRATION NUMBER:  40,403
;      REFERENCE/DOCKET NUMBER:  CGC 1695/CIP3/DIV6 - SQLv3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8582
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2004 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..2004
;      OTHER INFORMATION:  /note= "Maize optimized DNA
;      OTHER INFORMATION:  sequence for VIP1A(a) 80 kd protein from AB78"
US-08-471-044-18

  Query Match            2.2%;  Score 33.4;  DB 2;  Length 2004;
  Best Local Similarity   55.7%;  Pred. No. 2.2;
  Matches   64;  Conservative    0;  Mismatches   51;  Indels    0;  Gaps    0;

Qy        1375 CAAGAGAAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATC 1434
               || ||||||    |    ||| ||||  |||  ||   |  |   | ||||||||| |
Db        1357 CAGGAGAAGCTGAACAAGAACCGCGACTACTACATCAGCCTGTACATGAAGAGCGAGAAG 1416

Qy        1435 GACAAGCAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGA 1489
                |||  ||| | || |||| ||  | | |  |||  |    |  |  ||  ||||
Db        1417 AACACCCAGTGCGAGATCACCATCGACGGCGAGATATACCCCATCACCACCAAGA 1471


Search completed: October 22, 2003, 13:36:58
Job time : 133 secs
```

```
                          GenCore version 5.1.6
                Copyright (c) 1993 - 2003  Compugen Ltd.

 nucleic - nucleic search, using sw model

 1 on:          October 22, 2003, 07:24:21 ; Search time 426 Seconds
                                            (without alignments)
                                            9562.094 Million cell updates/sec

:le:            US-10-006-852-1
rfect score:    1509
quence:         1 atggtgctctcccacgccgt..........agacgagtggtatctgctaa 1509

oring table:    IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

arched:         2552756 seqs, 1349719017 residues

tal number of hits satisfying chosen parameters:      5105512

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
               Maximum Match 100%
               Listing first 45 summaries

itabase :       N_Geneseq_19Jun03:*
                1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
                25:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                                 SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                   Description
 ---------------------------------------------------------------------------
    1     1509  100.0   1509  24  AAL43410              A thaliana GAD1 co
    2   1507.4   99.9   1509  24  ABZ14005              Arabidopsis thalia
    3    883.4   58.5   1705  24  AAL43415              Tobacco GAD1 codin
    4    871.4   57.7   1784  24  AAL43417              Petunia GAD coding
    5    870.2   57.7   1665  24  AAL43411              A thaliana GAD2 co
    6    857.8   56.8   1771  24  AAL43416              Tobacco GAD2 codin
    7    854.2   56.6   1482  24  ABZ13132              Arabidopsis thalia
    8      770   51.0   1605  21  AAC36324              Arabidopsis thalia
    9    736.6   48.8   1783  15  AAQ74691              Early Ripening Tom
   10    736.6   48.8   1783  24  AAL43418              Tomato GAD coding
   11    711.4   47.1   1180  22  AAI70004              Tomato glutamate d
   12      402   26.6    416  25  ABX62354              Arabidopsis thalia
   13    388.6   25.8   1607  21  AAC44889              Arabidopsis thalia
   14    386.2   25.6   1545  22  AAH23436              Aspergillus oryzae
   15    385.8   25.6   2121  24  AAL43413              A thaliana GAD4 co
   16    360.6   23.9    742  19  AAV28662              Ripening banana pu
   17    351.8   23.3   1946  24  AAL43414              A thaliana GAD5 co
   18    330.4   21.9   2243  22  AAH23435              Aspergillus oryzae
   19    325.4   21.6    807  19  AAV28663              Ripening banana pu
c  20    311.2   20.6 4403765  22  AAI99683              Mycobacterium tube
c  21    311.2   20.6 4411529  22  AAI99682              Mycobacterium tube
   22      310   20.5   2493  24  AAL43412              A thaliana GAD3 co
c  23    271.8   18.0 2365589  24  ABA90521              Genomic sequence o
   24    270.6   17.9   5565  19  AAV11905              L. lactis NS3 locu
   25    259.4   17.2    440  24  ABL94157              Arabidopsis thalia
   26    256.2   17.0   2742  24  ABQ69218              Listeria monocytog
   27    256.2   17.0   6077  24  ABQ71043              Listeria monocytog
c  28    249.8   16.6 2944528  24  ABA03041              Listeria monocytog
   29    242.6   16.1    387  25  ABX19421              Human GDP-mannose
c  30    240.2   15.9 684707  24  ABQ67196              Listeria innocua c
c  31    240.2   15.9 3011208  24  ABQ69245              Listeria innocua D
   32    236.2   15.7   1389  24  ABQ69870              Listeria monocytog
   33    236.2   15.7 2944528  24  ABA03041              Listeria monocytog
   34    219.6   14.6 684707  24  ABQ67196              Listeria innocua c
   35    219.6   14.6 3011208  24  ABQ69245              Listeria innocua D
   36    206.8   13.7   4182  24  ABQ70867              Listeria monocytog
   37    187.2   12.4    280  25  ABX25243              Human GDP-mannose
   38      186   12.3   3129  23  AAS81862              DNA encoding novel
   39      186   12.3   3129  23  AAS89927              DNA encoding novel
c  40      186   12.3   3129  23  AAS92819              DNA encoding novel
   41    156.2   10.4    277  25  ABX31139              Human GDP-mannose
   42    143.2    9.5   1830  23  AAS81857              DNA encoding novel
   43    141.6    9.4    713  21  AAF13545              Aspergillus oryzae
c  44    136.6    9.1   3201  23  AAS93387              DNA encoding novel
   45    136.6    9.1   3468  23  AAS89926              DNA encoding novel

                                ALIGNMENTS

RESULT 1
AAL43410
ID   AAL43410 standard; cDNA; 1509 BP.
XX
AC   AAL43410;
XX
DT   25-SEP-2002  (first entry)
XX
DE   A thaliana GAD1 coding sequence.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1509
FT                   /*tag=  a
FT                   /product= "GAD1"
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
```

```
  WPI; 2002-490073/52.
  P-PSDB; AAO15132.

  Making transformed plants that selectively increase gamma-aminobutyric
  acid production, by incorporating a DNA construct with a polynucleotide
  encoding a plant glutamic acid decarboxylase enzyme into plant's genome
   -

  Claim 18; Page 53; 63pp; English.

  The present invention relates to a method of producing a transformed
  plant that selectively increases production of gamma-aminobutyric acid
  (GABA) in response to a signal, by incorporating into the plant's genome
  a DNA construct with a non-constitutive promoter operably linked to a
  polynucleotide encoding a functional plant glutamic acid decarboxylase
  (GAD), to provide a transformed plant that expresses the GAD coding
  sequence in response to a signal. Plants of this type have an enhanced
  ability to tolerate environmental or other stresses. The present sequence
  is the A. thaliana GAD1 coding sequence.

  Sequence 1509 BP; 433 A; 321 C; 385 G; 370 T; 0 other;

Query Match             100.0%;  Score 1509;  DB 24;  Length 1509;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1509;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y           1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b           1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

y          61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

y         121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

y         181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

Qy        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Qy       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

Qy       1501 ATCTGCTAA 1509
              |||||||||
Db       1501 ATCTGCTAA 1509


RESULT 2
ABZ14005
ID   ABZ14005 standard; DNA; 1509 BP.
XX
AC   ABZ14005;
XX
DT   21-JAN-2003  (first entry)
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 1810.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS   Arabidopsis thaliana.
XX
```

```
    WO200216655-A2.

    28-FEB-2002.

    24-AUG-2001; 2001WO-US26685.

    24-AUG-2000; 2000US-227866P.
    26-JAN-2001; 2001US-264647P.
    22-JUN-2001; 2001US-300111P.

    (SCRI ) SCRIPPS RES INST.
    (SYGN ) SYNGENTA PARTICIPATIONS AG.

I   Harper JF,  Kreps J,  Wang X,  Zhu T;

    WPI; 2002-304127/34.

    Identifying a stress condition to which a plant cell has been exposed
    and producing plants with increased tolerance to these abiotic stresses
    -
X
S   Claim 144; SEQ ID NO 1810; 577pp + Sequence Listing; English.
X
C   The invention relates to identifying a stress condition to which a plant
C   cell has been exposed, comprising:
C   (a) contacting nucleic acid representative of expressed polynucleotides
C   in the plant cell with an array or probes representative of the plant
C   cell genome; and
C   (b) detecting a profile of expressed polynucleotides in the plant cell
C   characteristic of a stress response. The method is useful in the
C   production of transgenic plants, cells and seeds and in producing plants
C   with increased tolerance to abiotic stress. The present sequence is that
C   of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
C   in methods of the invention.
C   Note: The sequence data for this patent is not represented in the printed
C   specification but is based on sequence information supplied to Derwent by
C   the European Patent Office.
X
Q   Sequence 1509 BP; 432 A; 322 C; 385 G; 370 T; 0 other;

 Query Match             99.9%;  Score 1507.4;  DB 24;  Length 1509;
 Best Local Similarity   99.9%;  Pred. No. 0;
 Matches 1508;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

Qy        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Qy       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

Qy       1501 ATCTGCTAA 1509
              |||||||||
```

```
)       1501 ATCTGCTAA 1509

SULT 3
L43415
)   AAL43415 standard; cDNA; 1705 BP.
K
C   AAL43415;
K
T   25-SEP-2002  (first entry)
K
E   Tobacco GAD1 coding sequence.
K
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
K
S   Nicotiana tabacum.
X
H   Key             Location/Qualifiers
T   CDS             71..1561
T                   /*tag=  a
T                   /product= "GAD1"
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I   Kinnersley AM,  Turano FJ;
X
R   WPI; 2002-490073/52.
R   P-PSDB; AAO15137.
X
T   Making transformed plants that selectively increase gamma-aminobutyric
T   acid production, by incorporating a DNA construct with a polynucleotide
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T   -
X
S   Claim 18; Page 56-57; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
C   (GABA) in response to a signal, by incorporating into the plant's genome
C   a DNA construct with a non-constitutive promoter operably linked to a
C   polynucleotide encoding a functional plant glutamic acid decarboxylase
C   (GAD), to provide a transformed plant that expresses the GAD coding
C   sequence in response to a signal. Plants of this type have an enhanced
C   ability to tolerate environmental or other stresses. The present sequence
C   is the tobacco GAD1 coding sequence.
X
Q   Sequence 1705 BP; 523 A; 301 C; 385 G; 496 T; 0 other;

  Query Match             58.5%;  Score 883.4;  DB 24;  Length 1705;
  Best Local Similarity   75.2%;  Pred. No. 1.5e-260;
  Matches 1125;  Conservative    0;  Mismatches  351;  Indels   21;  Gaps    1;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||| || ||| |  | |  |||||    ||||||||| |||||||||| ||||| || 
Db         71 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 130

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || || || |||||||| ||||| ||||| |||||||| || || ||||| || ||||||
Db        131 CGATATGTTCGTACTTCTCTTCCGAGGTTTAAGATGCCAGAGAATTCGATACCAAAGGAA 190

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              || || ||||| ||||| || || ||||| ||| | || || |||||| |  | || |||
Db        191 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGACTAAATTTA 250

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              || || |||||||| ||||||||||| || |||||| | ||||| || ||||    |||| 
Db        251 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAACAAACTGATGATGGATTCCATT 310

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              |||||||| || |||||||||||| || ||||| || ||||| ||||||||||| ||||||
Db        311 AACAAGAATTACGTTGACATGGATGAATACCCTGTAACCACTGAACTTCAGAATCGATGT 370

Qy        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              || ||||||||| || ||| | || || |||||  | | ||| |  ||||| || || |||
Db        371 GTAAACATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 430

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              || ||||| |||||||| || ||||| || ||| | || || || || ||||| | |||
Db        431 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 490

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||| || ||    |||||  ||||||| |||   || || ||||| |||||||| || 
Db        491 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTGTGACAAGCCCAATATTGTCACTGGT 550

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              |||||||| || ||||||||||||||||| || ||||| ||||| || || || ||||||
Db        551 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 610

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              || || |||||||| ||||||||||||||||||||| |  ||||||| || ||||| |||
Db        611 GTAAAGTTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 670

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              |||||||| |||||||| || |  ||  | |||||||| || ||||| ||||| ||||||
Db        671 GAGAACACAATTTGTGTAGCTGCTATCTTGGGTTCCACACTCAATGGTGAATTTGAAGAT 730

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| | |||||| || ||||| | || |||||||||||||||||| ||||| || |||
Db        731 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 790

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || || |||||||| ||||||||| ||||| |||||||| ||  | ||||| ||  | |||
Db        791 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 850

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              |||||||||||| | || |||||||||||||| || |||||||||||||| ||||| |||
Db        851 TGGGACTTTAGATTGCCATTGGTGAAGAGTATAAACGTGAGTGGTCACAAATATGGTCTT 910

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || || || || ||||||||||  || ||||| || || || || || ||||| ||||| 
Db        911 GTTTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTACCTGACGAACTT 970

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              |||||||| || ||||||||||||||||| ||||| || || |||||||| ||||| |||
Db        971 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1030

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              ||||| || ||||| |||||||||| |||||||||||| || |||||    |||||||||
Db       1031 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1090

Qy       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              |  ||||| |||||||||||   ||| |||   |  || ||||| |||||||||||| || 
Db       1091 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1150

Qy       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              |  | ||| ||||||||  | || || || |   ||||| ||| | || ||||||||  | 
Db       1151 AGTGGAAGATTCAACATAATATCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1210

Qy       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              ||||| | |||    |||| |||||||||||| || || |  ||| | || | |||||||
Db       1211 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1270

Qy       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              ||  | || || || || |||||| |||||||| |||||  |||| |||||  | || || 
Db       1271 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1330

Qy       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
```

```
           || ||||||||||||||  | |||||||| ||| |||| || || || || || |||||
      1331 ATTAGAGAAGATTTCTCCCGTACACTCGCCGAGCGACTGGTAATAGACATTGAAAAAGTC 1390

      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
            | |  ||||| ||    |||||  |||| || |     ||  || |  | |   | | |
      1391 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTGGCCGAGGCG 1450

      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
           ||  |                      |  | |  |   ||||| |  ||||  ||   |
)     1451 AATGG---------------------CAGCGGCGTGCATAAGAAAACAGATAGAGAAGTG 1489

/     1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
           |||   || || |  ||||  |||||||| ||||| || || | |||||||| || |
)     1490 CAGCTTGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGACT 1546

ESULT 4
AL43417
D   AAL43417 standard; DNA; 1784 BP.
X
C   AAL43417;
X
T   25-SEP-2002  (first entry)
X
E   Petunia GAD coding sequence.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
X
S   Petunia sp.
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I   Kinnersley AM,  Turano FJ;
X
R   WPI; 2002-490073/52.
R   P-PSDB; AAO15139.
X
T   Making transformed plants that selectively increase gamma-aminobutyric
T   acid production, by incorporating a DNA construct with a polynucleotide
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T    -
X
S   Claim 18; Page 58; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
C   (GABA) in response to a signal, by incorporating into the plant's genome
C   a DNA construct with a non-constitutive promoter operably linked to a
C   polynucleotide encoding a functional plant glutamic acid decarboxylase
C   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the petunia GAD coding sequence.
XX
SQ   Sequence 1784 BP; 550 A; 298 C; 400 G; 536 T; 0 other;

  Query Match             57.7%;  Score 871.4;  DB 24;  Length 1784;
  Best Local Similarity   74.9%;  Pred. No. 7.4e-257;
  Matches 1121;  Conservative    0;  Mismatches  366;  Indels   10;  Gaps    2;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||| || ||  |  | || ||| ||   || || ||| | |||||||| || || ||
Db         72 ATGGTTCTATCAAAGACAGTGTCGCAGAGCGATGTGTCCATTCACTCCACGTTTGCTTCT 131
```

```
Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || || || || ||||| ||||| ||||| || ||||| || || ||||| || || |||
Db        132 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGATAATTCGATACCAAAAGAA 191

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              || || ||||||||||| || || || |||||| | || || || ||| || |||||||
Db        192 GCAGCATATCAGATCATAAATGATGAACTGATGTTAGATGGAAACCCAAGGCTGAACTTG 251

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              || || ||||| || ||||||||||| || |||||||||||  | || |||  ||| ||
Db        252 GCTTCTTTTGTTACAACATGGATGGAACCAGAGTGTGATAAGTTGATGATGGACTCTATT 311

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||||||||| ||||| || || || || ||||| || |||||||| ||||||
Db        312 AACAAGAACTATGTTGATATGGATGAATATCCTGTTACCACTGAGCTTCAGAATCGATGT 371

Qy        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              || |||||||| || ||| | || ||||||||  | ||||| |  || || || || |||
Db        372 GTAAACATGATAGCTCATTTGTTTAATGCACCACTTGAAGATGGAGAAACTGCAGTTGGA 431

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              || ||||| |||||||| || || ||||| ||| | || || || || |||||| | |||
Db        432 GTTGGAACTGTTGGATCCTCTGAAGCCATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 491

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              |||||||||||    |||||  |||||||||||   || || ||||||||||| || ||
Db        492 TGGCAGAACAAAATGAAAGCCCAAGGCAAACCCTGTGACAAGCCCAACATTGTTACTGGT 551

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              || ||||| || ||||| ||||||||||| || ||||| ||||| || || || ||||||
Db        552 GCAAATGTCCAGGTGTGCTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 611

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              || ||  | ||||||||||||||||||||||||||| |  ||||||| || ||||| |||
Db        612 GTAAAGCTTAGTGAAGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAGATGGTGGAT 671

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              || |||||||||||||| || |  || |  |||||||| || ||||||||||| |||||
Db        672 GAAAACACCATTTGTGTAGCTGCTATCTTAGGTTCCACCCTCAATGGAGAATTTGAAGAC 731

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| | |||||| |||||||||||||| ||||||||||||||||| ||||| || |||
Db        732 GTTAAGCGCTTGAATGATCTCTTGGTCGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 791

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || || |||||||| |||||||| ||||| |||||||||||  | || || ||  | ||
Db        792 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCGTTCATTTACCCAGAGCTTGAG 851

Qy        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              |||||||||||| | || || |||||||| || ||||| ||||||||||| ||||| |||
Db        852 TGGGACTTTAGATTGCCATTAGTGAAGAGCATTAATGTAAGTGGTCACAAATATGGTCTT 911

Qy        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || || || || |||||||||||  | ||||| ||| |||| ||||||||||| |||||
Db        912 GTCTATGCTGGTATTGGTTGGGTCGTTTGGAGGAAC-AAGATGATTTGCCTGATGAACTT 970

Qy        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              |||||||| || ||||||||||||||||| ||||| || || |||||||| || || |||
Db        971 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTTTCTAAA 1030

Qy        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              ||||| || ||||| ||||||||||| ||||||||||| || |||||  | |||||||||
Db       1031 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTATGAGGGTTAC 1090

Qy       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
              |  |||||||||||||||||   ||| |||      || ||||| ||||| || || |||
Db       1091 AAGAATGTGATGGAGAATTGTCAAGAAAATGCATCGGTACTAAGAGAAGGGCTAGAAAAG 1150

Qy       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              |||| ||| ||||||||  |||| || ||    ||||| ||  | || || |||||  |
Db       1151 ACAGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTACCTTTAGTAGCATTCTCTCTT 1210

Qy       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
```

```
              ||||| | |||    |||| |||||||| || || || |   |  | || | ||| |||
b        1211 AAAGACAACAGGCAACACAACGAGTTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1270

y        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              || || || || || || ||| ||||| |||||||||||| || |||||  | || ||
b        1271 ATTGTTCCTGCATATACTATGCCACCAAACGCACAACACATTACAGTTCTCAGAGTTGTG 1330

y        1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              |||||||||||||||||  | || || || ||  |||| || |  || || || ||||| 
b        1331 ATCAGAGAAGATTTCTCCCGTACGCTTGCAGAACGACTGGTAAGAGACATCGAAAAAGTC 1390

y        1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
               | | ||| || ||    || ||| |  | || | |   ||  |  |  |  |   | |||
b        1391 CTTCATGAACTTGACACACTCCCTGCACGTGTCAATGCTAAGCTCGCTGTGGCCGAGGAG 1450

y        1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
               ||  |    | ||  ||         |  |    |   ||||| |  |||| |||   |
b        1451 CAGGCGGCTGCGAATGGC---------AGCGAGGTGCATAAGAAAACAGATAGCGAAGTG 1501

y        1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
              |||   || || || ||||  ||||||||||||||| |  || | |||||||| || |
b        1502 CAGTTGGAGATGATAACTGCATGGAAGAAGTTTGTTGAAGAAAAGAAGAAGAAGACT 1558

ESULT 5
AL43411
D   AAL43411 standard; cDNA; 1665 BP.
X
C   AAL43411;
X
T   25-SEP-2002  (first entry)
X
E   A thaliana GAD2 coding sequence.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
X
S   Arabidopsis thaliana.
X
H   Key             Location/Qualifiers
T   CDS             17..1501
T                   /*tag=  a
T                   /product= "GAD2"
X
N   WO200238736-A2.
.X
'D  16-MAY-2002.
(X
'F  07-NOV-2001; 2001WO-US47447.
(X
'R  07-NOV-2000; 2000US-246367P.
(X
'A  (EMER-) EMERALD BIOAGRICULTURE CORP.
(X
'I  Kinnersley AM,  Turano FJ;
(X
)R  WPI; 2002-490073/52.
)R  P-PSDB; AAO15133.
(X
PT  Making transformed plants that selectively increase gamma-aminobutyric
PT  acid production, by incorporating a DNA construct with a polynucleotide
PT  encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT  -
XX
PS  Claim 18; Page 53-54; 63pp; English.
XX
CC  The present invention relates to a method of producing a transformed
CC  plant that selectively increases production of gamma-aminobutyric acid
CC  (GABA) in response to a signal, by incorporating into the plant's genome
CC  a DNA construct with a non-constitutive promoter operably linked to a
CC  polynucleotide encoding a functional plant glutamic acid decarboxylase
CC  (GAD), to provide a transformed plant that expresses the GAD coding
CC  sequence in response to a signal. Plants of this type have an enhanced
CC  ability to tolerate environmental or other stresses. The present sequence
CC  is the A. thaliana GAD2 coding sequence.
XX
SQ  Sequence 1665 BP; 490 A; 311 C; 423 G; 441 T; 0 other;

  Query Match             57.7%;  Score 870.2;  DB 24;  Length 1665;
  Best Local Similarity   75.1%;  Pred. No. 1.7e-256;
  Matches 1110;  Conservative    0;  Mismatches  348;  Indels   21;  Gaps    1;

Qy         31 GACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCACTTCCTAGGTTC 90
              || |  || |||  | |||  |||| ||| || || || || ||| ||||||| | ||
Db         44 GATGAATCTGTCTGCACCATGTTCGGATCTCGCTATGTTCGCACTACACTTCCCAAGTAT 103

Qy         91 AAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTG 150
               ||||    || || ||||| || || || || || |||||||||| ||  || ||||||
Db        104 GAGATTGGTGAGAATTCGATACCGAAAGACGCTGCATATCAGATCATAAAAGATGAGCTG 163

Qy        151 ATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCT 210
              |||||||| || || ||  || | ||| |||| || |||||||| |||||||||||| ||
Db        164 ATGCTTGATGGTAACCCGAGGCTTAACCTAGCTTCGTTTGTGACTACATGGATGGAACCA 223

Qy        211 GAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACATGGACGAGTAC 270
              |||||||| |||||||||||  ||| ||||||||||||||| ||||| ||||| ||||||
Db        224 GAGTGTGACAAACTCATCATGGACTCTATCAACAAGAACTACGTTGATATGGATGAGTAC 283

Qy        271 CCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATCTATTCAATGCA 330
              || |||| ||| || || ||||||||||||||| ||||| || || |  || |||||||| 
Db        284 CCTGTCACAACTGAGCTCCAGAACCGATGTGTAAACATTATAGCTCGACTGTTCAATGCG 343

Qy        331 CCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATA 390
              ||  | || ||  | ||||| || || |||||||| || ||||| || || || ||||| 
Db        344 CCACTCGAGGAATCTGAGACGGCGGTGGGAGTAGGGACAGTTGGTTCTTCAGAAGCCATC 403

Qy        391 ATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAA 450
              ||||| ||||| |||||||||||  | ||||||||||||| ||||| ||||| || |||
Db        404 ATGTTAGCCGGATTGGCCTTCAAAAGAAAATGGCAGAACAAACGCAAGGCTGAGGGTAAA 463

Qy        451 CCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTC 510
              |||   || |||||||||||||||| |||||||||||||||| || ||||||||||||
Db        464 CCCTATGACAAACCCAACATTGTCACTGGAGCCAATGTTCAAGTTTGCTGGGAGAAATTC 523

Qy        511 GCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATG 570
              ||| ||||||| ||||| || || |||||||| ||  | ||||||||| |||||| ||||||
Db        524 GCTCGGTACTTCGAGGTGGAGCTAAAGGAAGTAAACCTAAGTGAAGGTTACTACGTGATG 583

Qy        571 GACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTT 630
              || ||  |  |||| |  || ||||| || |||||||| || ||||| || | |||  |
Db        584 GATCCAGACAAAGCAGCAGAAATGGTAGACGAGAACACAATCTGTGTCGCAGCCATATTG 643

Qy        631 GGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAA 690
              || ||||| || || || || |||||||| || ||||   | || ||  |  | ||| |
Db        644 GGATCCACACTCAACGGTGAGTTCGAAGACGTGAAACGTCTCAATGACTTGCTAGTCAAG 703

Qy        691 AAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTC 750
              || ||| | || || || || | ||||| |||||||||||||||| |||||||||||| |||
Db        704 AAAAACGAGGAGACTGGTTGGAACACACCGATCCACGTGGATGCAGCAAGTGGAGGGTTC 763

Qy        751 ATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGT 810
              || || |||||| | ||||| ||||| ||||||||||||||||||||||| ||||| ||||||
Db        764 ATAGCTCCGTTTATCTATCCTGAATTAGAATGGGACTTTAGACTTCCTTTGGTTAAGAGT 823

Qy        811 ATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGG 870
              ||||| ||||||||||||||||||||||| || || || || ||||||||||||  | |||
Db        824 ATCAACGTGAGTGGTCACAAGTATGGACTGGTCTATGCTGGTATTGGTTGGGTCGTGTGG 883

Qy        871 AGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGAC 930
              ||      ||||||||||||||| || || ||||| ||||| | ||||||||||||||||| 
Db        884 AGGGCAGCAGAGGATTTGCCTGAAGAGCTTATCTTTCATATTAATTATCTTGGTGCTGAT 943

Qy        931 CAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTAC 990
              |||||||| || ||||||||||||||||||| || || || || | ||||||||||||||||
Db        944 CAACCCACTTTCACTCTCAATTTCTCCAAGGGATCGAGCCAAATTATTGCTCAATACTAC 1003
```

```
          991 CAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAAT 1050
              || || || ||  | ||   |||||| |||| |||||||||||||||||||| ||||||
         1004 CAGCTCATTCGTCTTGGATTCGAGGGGTACAAAAATGTGATGGAGAATTGCATAGAGAAC 1063

'        1051 ATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGAC 1110
              ||| | || || |  |||||  | ||||| |||||  | |||||||| ||||||||||||
)        1064 ATGGTGGTTCTCAAAGAAGGGATAGAGAAAACAGAGCGTTTCAACATAGTCTCAAAGGAC 1123

'        1111 GAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAA 1170
               | ||||||||| | || || |||||  | || ||    || |  ||||  ||||||||
)        1124 CAAGGAGTGCCAGTCGTAGCCTTCTCTCTCAAGGACCATAGTTTCCACAACGAGTTCGAG 1183

y        1171 ATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAAT 1230
              ||||| || ||||| ||  | | ||| ||||| || || || ||||| ||||||| |  ||
)        1184 ATCTCTGAGATGCTACGTCGTTTTGGCTGGATCGTCCCAGCTTACACTATGCCTGCCGAT 1243

y        1231 GCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCT 1290
              ||||| ||||||||| ||||| ||||| || ||||| ||||||||||| |||||||||||
)        1244 GCACAGCACATCACGGTTCTGCGTGTTGTCATCAGGGAAGATTTCTCAAGAACACTCGCG 1303

y        1291 GAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGCTTCCTTCGAGA 1350
              |||||||||||    |||||   ||| ||| | | |||||| |||    | ||||| |
)        1304 GAGAGACTTGTTGCTGATATTTCGAAGGTGCTTCATGAGCTAGATACCTTGCCTTCCAAG 1363

y        1351 GTGATTCACAAAATATCACTTGGACAAGAGAAGAGTGAATCTAACAGCGATAACTTGATG 1410
               |                     |    ||||||  | |    |  |          |
b        1364 AT---------------------ATCTAAGAAGATGGGAATAGAAGGGATCGCGGAAAAT 1402

y        1411 GTCACGGTGAAGAAGAGCGATATCGACAAGCAGAGAGATATCATCACTGGCTGGAAGAAG 1470
              || | || ||||||||  || |  || |  | ||  ||  | ||   ||| |||| ||||
b        1403 GTAAAGGAGAAGAAGATGGAGAAGGAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAG 1462

y        1471 TTTGTCGCCGACAGGAAGAAGACGAGTGGTATCTGCTAA 1509
              |||||    || |||||||||| || |||| | ||||||
b        1463 TTTGTGAAGGAGAGGAAGAAGATGAATGGTGTGTGCTAA 1501

ESULT 6
AL43416
D   AAL43416 standard; cDNA; 1771 BP.
X
C   AAL43416;
X
T   25-SEP-2002  (first entry)
X
E   Tobacco GAD2 coding sequence.
X
.W  GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
X
)S  Nicotiana tabacum.
X
H   Key             Location/Qualifiers
T   CDS             67..1557
T                   /*tag=  a
T                   /product= "GAD2"
X
N   WO200238736-A2.
X
PD  16-MAY-2002.
XX
PF  07-NOV-2001; 2001WO-US47447.
XX
PR  07-NOV-2000; 2000US-246367P.
XX
PA  (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI  Kinnersley AM,  Turano FJ;
XX
DR  WPI; 2002-490073/52.
DR  P-PSDB; AAO15138.
XX
PT  Making transformed plants that selectively increase gamma-aminobutyric
PT  acid production, by incorporating a DNA construct with a polynucleotide
PT  encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT   -
XX
PS  Claim 18; Page 57; 63pp; English.
XX
CC  The present invention relates to a method of producing a transformed
CC  plant that selectively increases production of gamma-aminobutyric acid
CC  (GABA) in response to a signal, by incorporating into the plant's genome
CC  a DNA construct with a non-constitutive promoter operably linked to a
CC  polynucleotide encoding a functional plant glutamic acid decarboxylase
CC  (GAD), to provide a transformed plant that expresses the GAD coding
CC  sequence in response to a signal. Plants of this type have an enhanced
CC  ability to tolerate environmental or other stresses. The present sequence
CC  is the tobacco GAD2 coding sequence.
XX
SQ  Sequence 1771 BP; 548 A; 295 C; 393 G; 535 T; 0 other;

  Query Match             56.8%;  Score 857.8;  DB 24;  Length 1771;
  Best Local Similarity   74.1%;  Pred. No. 1.1e-252;
  Matches 1109;  Conservative    0;  Mismatches  367;  Indels   21;  Gaps    1;

Qy          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
              ||||| || ||| |  | |  |||||    ||||||||||| |||||||| ||||| ||
Db         67 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCGTTCACTCCACTTTCGCCTCC 126

Qy         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
              || || || || ||||| ||||| ||||| || ||||| || || || || || ||||||
Db        127 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGAGAATTCAATACCAAAGGAA 186

Qy        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
              || || |||||||| || || || ||||| ||| | || || |||||| || | || |||
Db        187 GCAGCATATCAGATTATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 246

Qy        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
              || || || || || |||||||||||||| || ||| |||   | || |||   |||||
Db        247 GCATCTTTCGTTACAACATGGATGGAGCCAGAATGTAATACGTTAATGATGGATTCCATT 306

Qy        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
              ||||||||||| ||||||||||| || ||||| || ||||| || |||||||| ||||||
Db        307 AACAAGAACTACGTTGACATGGATGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 366

Qy        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
              || || ||||| || ||| | || ||||||||  | | ||| |  ||||| || || |||
Db        367 GTAAATATGATAGCTCATTTGTTTAATGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 426

Qy        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
              || ||||| |||||||| || || || || ||| ||   || || ||||| ||| | |||
Db        427 GTTGGAACTGTTGGATCCTCTGAAGCTATTATGCTTGCTGGATTAGCCTTTAAGAGAAAA 486

Qy        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
              ||||| || ||    |||||  ||||||| ||| | ||||| ||||| |||||||||||
Db        487 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTTTGATAAGCCCAATATTGTCACCGGT 546

Qy        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
              || ||||| || ||||||||||||||||| || ||||| ||||| || ||  | || |||
Db        547 GCTAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 606

Qy        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
              || ||||||||||| ||||||||||||||||||||| |  ||||||| || ||||| |||
Db        607 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 666

Qy        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              ||||| |||||||||||||| |  ||  | ||||| || || ||||| ||||| ||||||
Db        667 GAGAATACCATTTGTGTTGCTGCTATCTTAGGTTCAACACTCAATGGTGAATTTGAAGAT 726

Qy        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| |  ||||| || || ||| | || ||||||||||||||||| ||||| || |||
Db        727 GTTAAGCGTTTGAATGACCTTTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 786

Qy        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || || |||||||| |||||||| ||||| |||||||| ||  | ||||| ||  | |||
Db        787 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 846
```

```
y        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
             |||||||||||| | || |||| ||||||||| ||||||||||||||||| ||||| |||
b        847 TGGGACTTTAGATTGCCATTGGAGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 906

y        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
             || || || || ||||||||||  || ||||| || || || || |||||||| |||||
b        907 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 966

y        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
             || ||||| ||||||||| ||||||||||| ||||| || || |||||||| ||||| |||
b        967 ATTTTCCACATCAATTACCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1026

y        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
             ||||| || ||||| |||||||||||| |||||||||||| || |||||    |||||||||
b       1027 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1086

y       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
             |  ||||| |||||||||||   ||| |||   |  ||  |||| ||||| | |||| ||
b       1087 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTATTAAGAGAAGGAATTGAAAAA 1146

y       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
             |  | ||| ||||||||  |||| || ||    ||||| ||  | || || || ||  |
b       1147 AGTGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTTCCCTTAGTAGCATTTTCTCTT 1206

y       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
             ||||| | |||    |||| |||||||||||| || || |  ||| | || | |||||||
b       1207 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1266

y       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
             || || | ||| || || ||||| |||||||| |||||  |||| |||||  | || ||
b       1267 ATTGTTCTGGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1326

y       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
             || ||||||||||||||  | ||||| || ||| |||| || || || || || |||||
b       1327 ATTAGAGAAGATTTCTCCCGCACACTAGCGGAGCGACTGGTAATAGACATTGAAAAAGTC 1386

y       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
              | |  |   | ||  | |||||  |||| || |     ||  || |  |  |   | | |
b       1387 TTCCACGGAGTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTGGCCGAGGCG 1446

y       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440
             ||  |                      |  | |  |   ||||| |  ||||  ||   |
b       1447 AATGG---------------------CAGCGGCGTGCATAAGAAAACAGATAGAGAAGTG 1485

y       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGT 1497
             |||  ||| || |  ||||  |||  ||| ||||| || || | |||||||| || |
b       1486 CAGCTAGAGATTACTACTGCATGGTTGAAATTTGTTGCTGATAAGAAGAAGAAGACT 1542

ESULT 7
BZ13132
D   ABZ13132 standard; DNA; 1482 BP.
X
C   ABZ13132;
X
T   21-JAN-2003  (first entry)
X
E   Arabidopsis thaliana stress regulated gene SEQ ID NO 937.
X
W   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS  Arabidopsis thaliana.
XX
PN  WO200216655-A2.
XX
PD  28-FEB-2002.
XX
PF  24-AUG-2001; 2001WO-US26685.
XX
PR  24-AUG-2000; 2000US-227866P.
PR  26-JAN-2001; 2001US-264647P.
PR  22-JUN-2001; 2001US-300111P.
```

```
XX
PA  (SCRI ) SCRIPPS RES INST.
PA  (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI  Harper JF,  Kreps J,  Wang X,  Zhu T;
XX
DR  WPI; 2002-304127/34.
XX
PT  Identifying a stress condition to which a plant cell has been exposed
PT  and producing plants with increased tolerance to these abiotic stresses
PT  -
XX
PS  Claim 144; SEQ ID NO 937; 577pp + Sequence Listing; English.
XX
CC  The invention relates to identifying a stress condition to which a plant
CC  cell has been exposed, comprising:
CC  (a) contacting nucleic acid representative of expressed polynucleotides
CC  in the plant cell with an array or probes representative of the plant
CC  cell genome; and
CC  (b) detecting a profile of expressed polynucleotides in the plant cell
CC  characteristic of a stress response. The method is useful in the
CC  production of transgenic plants, cells and seeds and in producing plants
CC  with increased tolerance to abiotic stress. The present sequence is that
CC  of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC  in methods of the invention.
CC  Note: The sequence data for this patent is not represented in the printed
CC  specification but is based on sequence information supplied to Derwent by
CC  the European Patent Office.
XX
SQ  Sequence 1482 BP; 402 A; 331 C; 378 G; 371 T; 0 other;

  Query Match             56.6%;  Score 854.2;  DB 24;  Length 1482;
  Best Local Similarity   76.0%;  Pred. No. 1.3e-251;
  Matches 1054;  Conservative    0;  Mismatches  333;  Indels    0;  Gaps    0;

Qy         1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
             |||||  | ||  |  | || || || || || |||||  |||| || || || || ||
Db         1 ATGGTTTTGTCTAAGACAGTTTCCGAATCTGATGTCTCAATCCATTCAACTTTTGCTTCT 60

Qy        61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
             ||||||||||| |  || |||||  | ||| | ||||| || ||||| || || || |||
Db        61 CGTTACGTCCGCAACTCTCTTCCACGATTCGAAATGCCTGAGAACTCAATCCCAAAAGAA 120

Qy       121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
             || || || || ||||||||||||||||| ||||| || || || ||| || ||||| ||
Db       121 GCAGCTTACCAAATCATCAACGACGAGCTAATGCTCGATGGTAACCCAAGGCTGAACCTA 180

Qy       181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
             || ||||| ||||| |||||||||||||| || ||||| || ||||| |||   ||||||
Db       181 GCTTCCTTCGTGACCACATGGATGGAGCCAGAATGTGACAAGCTCATGATGGAGTCCATC 240

Qy       241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
             ||||||||||| || ||||||||||||||||| ||||||||| || |||||||||||||||
Db       241 AACAAGAACTACGTCGACATGGACGAGTACCCTGTCACCACTGAGCTTCAGAACCGATGT 300

Qy       301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
             || |||||||| |||| ||| ||||| || ||| | |  || |  ||  | ||||| ||
Db       301 GTTAACATGATAGCACGTCTCTTCAACGCGCCGCTTGGTGACGGTGAAGCTGCCGTTGGT 360

Qy       361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420
             || || ||||| ||||| |||||||| || ||||||||||||||||| || ||| |  ||
Db       361 GTTGGCACCGTCGGATCGTCGGAGGCGATTATGTTGGCCGGTTTGGCTTTTAAGAGACAA 420

Qy       421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480
             |||||||| ||||| || ||  ||||    ||    ||||| ||||| || || |||||
Db       421 TGGCAGAATAAGCGTAAGGCCCAAGGGCTTCCTTATGATAAGCCCAATATCGTAACCGGT 480

Qy       481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540
             || ||||| || || || |||||||||||||| ||||| || || || || |||||||||
Db       481 GCTAATGTCCAGGTTTGCTGGGAGAAATTCGCAAGGTATTTCGAAGTGGAGCTTAAGGAA 540

Qy       541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
             |||||  | || ||||  || || ||||||||||||  | | || || || ||||| ||
```

```
)         541 GTGAACCTAAGAGAAGACTATTACGTGATGGACCCTGTAAAGGCGGTCGAAATGGTAGAC 600

/         601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660
              || ||||| |||||||| || | ||| || ||||| ||  | |  || |||||||||||
)         601 GAAAACACAATTTGTGTCGCTGCCATCCTCGGTTCAACGTTAACCGGTGAATTCGAAGAC 660

/         661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720
              ||||| ||| | ||||| ||| | ||||| || |||||  ||||||||||||| || |||
)         661 GTTAAGCTCCTCAACGACCTCCTTGTCGAGAAAAACAAGCAAACCGGATGGGACACGCCA 720

y         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780
              || |||||||| || || ||||| || || ||||| ||||| |||||||||||  ||||
b         721 ATACACGTGGACGCAGCGAGTGGTGGGTTTATTGCTCCGTTCTTGTATCCGGAGCTGGAG 780

y         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840
              ||||| ||  | || || ||||| |||||||| ||||||||||||||||| || ||  |
b         781 TGGGATTTCCGGCTACCGTTGGTTAAGAGTATTAATGTGAGTGGTCACAAATACGGTTTG 840

y         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900
              || ||||| || |||||||||||  | ||||||| ||||   ||||||||||| |||||
b         841 GTTTACGCCGGTATTGGTTGGGTTGTATGGAGAACCAAAACCGATTTGCCTGATGAACTT 900

y         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960
              ||||||||||||||||||||||| ||||| ||||| |||||||| ||||| |||||||||
b         901 ATCTTCCATATCAATTATCTTGGCGCTGATCAACCAACCTTTACACTCAACTTCTCCAAA 960

y         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020
              |||||||||||||| |||||||| |||||||| || || ||  | |    |||||||||
b         961 GGTTCAAGTCAAGTGATTGCTCAGTACTACCAGCTGATTCGTCTTGGATTCGAGGGTTAT 1020

y        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080
               | ||||||||||| |||||  | || || ||||| || |||||  ||||| | |||||
b        1021 CGCAATGTGATGGATAATTGTCGGGAAAACATGATGGTACTAAGACAAGGATTAGAGAAA 1080

y        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140
              || | | | || || |||||||| || ||  | || || ||  | || || || ||  |
b        1081 ACGGGACGTTTTAAAATCGTCTCCAAAGAAAACGGTGTTCCGTTAGTGGCGTTTTCTCTC 1140

y        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
              |||||||| ||| | ||||  ||||||||  |  || | |  || ||  | |  || |||
b        1141 AAAGATAGTAGCCGCCACAACGAGTTCGAGGTGGCCCATACACTCCGTCGCTTCGGCTGG 1200

y        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
              || || ||||||||||| |||||| |  |||| || |   ||||||| ||||| || |||
b        1201 ATCGTTCCGGCCTACACGATGCCTGCGGATGCGCAGCATGTCACTGTCCTTCGAGTTGTT 1260

y        1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320
              ||| |||||||||||||  ||||  | || |||||| | ||    ||| | ||||| ||
b        1261 ATCCGAGAAGATTTCTCTCGAACCTTAGCCGAGAGATTGGTAGCTGATTTCGAGAAGGTT 1320

)y       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380
               | |  |||||||||  ||||||  |||| ||      ||| ||  |   |||| |||
)b       1321 CTACACGAGCTCGATACGCTTCCGGCGAGGGTTCACGCCAAGATGGCTAATGGAAAAGTT 1380

)y       1381 AAGAGTG 1387
              ||  |||
)b       1381 AACGGTG 1387

RESULT 8
AAC36324
ID   AAC36324 standard; DNA; 1605 BP.
XX
AC   AAC36324;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 13376.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
```

```
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
```

```
     30-JUN-1999;   99US-0141287.
     01-JUL-1999;   99US-0141842.
     01-JUL-1999;   99US-0142154.
     02-JUL-1999;   99US-0142055.
     06-JUL-1999;   99US-0142390.
     08-JUL-1999;   99US-0142803.
     09-JUL-1999;   99US-0142920.
     12-JUL-1999;   99US-0142977.
     13-JUL-1999;   99US-0143542.
     14-JUL-1999;   99US-0143624.
     15-JUL-1999;   99US-0144005.
     16-JUL-1999;   99US-0144085.
     16-JUL-1999;   99US-0144086.
     19-JUL-1999;   99US-0144325.
     19-JUL-1999;   99US-0144331.
R    19-JUL-1999;   99US-0144332.
R    19-JUL-1999;   99US-0144333.
R    19-JUL-1999;   99US-0144334.
R    19-JUL-1999;   99US-0144335.
R    20-JUL-1999;   99US-0144352.
R    20-JUL-1999;   99US-0144632.
R    20-JUL-1999;   99US-0144884.
R    21-JUL-1999;   99US-0144814.
R    21-JUL-1999;   99US-0145086.
R    21-JUL-1999;   99US-0145088.
R    22-JUL-1999;   99US-0145085.
R    22-JUL-1999;   99US-0145087.
R    22-JUL-1999;   99US-0145089.
R    22-JUL-1999;   99US-0145192.
R    23-JUL-1999;   99US-0145145.
R    23-JUL-1999;   99US-0145218.
R    23-JUL-1999;   99US-0145224.
R    26-JUL-1999;   99US-0145276.
R    27-JUL-1999;   99US-0145913.
R    27-JUL-1999;   99US-0145918.
R    27-JUL-1999;   99US-0145919.
R    28-JUL-1999;   99US-0145951.
R    02-AUG-1999;   99US-0146386.
R    02-AUG-1999;   99US-0146388.
R    02-AUG-1999;   99US-0146389.
R    03-AUG-1999;   99US-0147038.
R    04-AUG-1999;   99US-0147204.
R    04-AUG-1999;   99US-0147302.
R    05-AUG-1999;   99US-0147192.
R    05-AUG-1999;   99US-0147260.
R    06-AUG-1999;   99US-0147303.
R    06-AUG-1999;   99US-0147416.
R    09-AUG-1999;   99US-0147493.
R    09-AUG-1999;   99US-0147935.
R    10-AUG-1999;   99US-0148171.
R    11-AUG-1999;   99US-0148319.
R    12-AUG-1999;   99US-0148341.
R    13-AUG-1999;   99US-0148565.
R    13-AUG-1999;   99US-0148684.
R    16-AUG-1999;   99US-0149368.
R    17-AUG-1999;   99US-0149175.
R    18-AUG-1999;   99US-0149426.
R    20-AUG-1999;   99US-0149722.
R    20-AUG-1999;   99US-0149723.
R    20-AUG-1999;   99US-0149929.
R    23-AUG-1999;   99US-0149902.
R    23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
```

```
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             51.0%;  Score 770;  DB 21;  Length 1605;
  Best Local Similarity   72.0%;  Pred. No. 1e-225;
  Matches 1004;  Conservative    0;  Mismatches  390;  Indels    0;  Gaps    0;

Qy         17 CCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTT 76
              ||   || || || ||||      | || ||||| || || ||  | || |||||| ||
Db         53 CCAACTCTGACTCCGACGAGCATTTGCATTCCACTTTTGCTTCTAGATATGTCCGTGCTG 112

Qy         77 CACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCA 136
                 |||| |||||||||||||| ||  | |  || || || || || || |||||  | |
Db        113 TTGTTCCCAGGTTCAAGATGCCTGACCATTGCATGCCCAAAGATGCTGCTTATCAAGTGA 172

Qy        137 TCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGA 196
              |||| || ||| ||||||||| || ||||||  || | ||| ||||||||||||| || |
Db        173 TCAATGATGAGTTGATGCTTGATGGTAATCCCAGGCTTAACCTAGCCTCCTTTGTCACCA 232

Qy        197 CATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTG 256
              | ||||||||| |||||||||| ||||||||||||   ||  |||| |||||||||||||
Db        233 CTTGGATGGAACCTGAGTGTGACAAACTCATCATGGATTCTGTCAATAAGAACTATGTTG 292

Qy        257 ACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCAC 316
              | ||||| || || || |||||||| || || |||||||| ||||| || ||||| |||
Db        293 ATATGGATGAATATCCTGTCACCACTGAGCTCCAGAACCGGTGTGTAAATATGATAGCAA 352

Qy        317 ATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGAT 376
              |  | ||| |||| ||  | | ||| |  ||| | ||  | ||    |||||| ||||| |
Db        353 ACTTTTTCCATGCTCCCGTTGGAGAAGACGAGGCTGCTATTGGGTGTGGAACTGTTGGTT 412

Qy        377 CATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCA 436
              |||| ||||| |||||| | || |||||||| |||||  | |||||||||  | | | | |
```

```
           413 CATCTGAGGCTATAATGCTTGCTGGTTTGGCTTTCAAAAGGAAATGGCAACATAGGAGAA 472

           437 AAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGT 496
               ||||| | ||   |||  | ||||| || |||||||||||| |||||||||||||| ||||
           473 AAGCTCAGGGTCTACCTATTGATAAGCCTAACATTGTCACTGGAGCCAATGTTCAGGTGT 532

r          497 GTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAG 556
               | |||||||| || || |||||||||||||| || || || || |||||| | |||||||
)          533 GCTGGGAGAAGTTTGCAAGGTACTTTGAGGTAGAGCTCAAAGAGGTGAAACTAAGTGAAG 592

r          557 GATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTG 616
                 |||||||| ||||| ||     ||||||| || |||||| |||||||| ||||| ||||
)          593 ACTACTATGTTATGGATCCAGCTAAAGCTGTAGAGATGGTGGATGAGAATACCATCTGTG 652

r          617 TTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACG 676
               |||| |  ||||| || || || |||| |||||| || || || ||||| |  |||||||
)          653 TTGCAGCAATTCTAGGATCTACACTTACTGGAGAGTTTGAGGACGTTAAGCAATTGAACG 712

r          677 ATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGG 736
               ||||||| |  || || |||  ||| || |||||||| || || || || || ||||| |
)          713 ATCTCTTAGCTGAGAAAAACGCAGAGACAGGATGGGAAACTCCTATTCATGTTGATGCAG 772

r          737 CAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTC 796
               | ||||||||||||||||| || ||  | || || ||  | |||||||||||||| ||||
)          773 CCAGTGGAGGATTCATTGCTCCTTTCCTCTACCCTGATCTTGAATGGGACTTTAGGCTTC 832

y          797 CCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTG 856
               | | |||||||||||| || || ||||||||||||||||||||||||||| |||||  |||
b          833 CATGGGTGAAGAGTATTAACGTCAGTGGTCACAAGTATGGACTTGTGTATGCAGGAGTTG 892

y          857 GTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATT 916
               |||||||  |||||||||  ||||| |||||||| ||||||||  ||||||| ||||| |
b          893 GTTGGGTTGTCTGGAGAACAAAAGATGATTTGCCAGAGGAACTTGTCTTCCACATCAACT 952

y          917 ATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCA 976
               |  | |  |||||| |||||||| || ||||||||| ||||| ||||| || |||| |||
b          953 ACTTGGGAGCTGATCAACCCACTTTCACTCTCAACTTCTCAAAAGGGTCGAGCCAAATCA 1012

y          977 TTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGA 1036
               ||||||| ||||| ||  |||||||| | |||   ||||| ||||  ||  | ||||| |
b         1013 TTGCTCAGTACTATCAGTTTATCCGACTAGGCTTTGAGGGATACAAGAACATAATGGAAA 1072

y         1037 ATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACA 1096
               | ||||  || ||    |     ||||| ||||||| | ||||| |||||  | ||||||||
b         1073 ACTGCATGGATAACGCAAGGAGGCTAAGAGAAGGAATAGAGATGACAGGGAAGTTCAACA 1132

y         1097 TCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTC 1156
               | || || || ||    || |||||||| || || |||||  | ||||| || |||   |
b         1133 TTGTGTCCAAAGATATTGGCGTGCCACTAGTGGCATTCTCTCTCAAAGACAGTAGCAAGC 1192

y         1157 ACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACA 1216
               |||| | ||| || ||| | ||     |  | |  |  || |||||  | ||||| ||||
b         1193 ACACGGTGTTTGAGATCGCAGAGTCTTTGAGAAAATTCGGGTGGATCATACCGGCTTACA 1252

)y        1217 CAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCT 1276
               | |||||| || ||||||| |||||  |||| ||  | || || || |||||||| ||
)b        1253 CTATGCCTGCAGATGCACAGCACATTGCTGTGCTCAGAGTTGTGATAAGAGAAGACTTTA 1312

)y        1277 CGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATG 1336
                  ||   || || || |||||  | |   ||||     | ||| ||   ||| | || |
)b        1313 GCCGAGGCCTTGCAGATAGACTCATCACACATATCATTCAGGTGCTGAAAGAGATTGAAG 1372

Qy        1337 AGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAGAAGAGTGAATCTAACA 1396
                |||||||   ||  |    ||        |   || |   |  |   ||||   | |
Db        1373 GGCTTCCTAGCAGGATTGCACATCTTGCTGCGGCTGCAGCGGTTAGTGGTGATGATGAAG 1432

Qy        1397 GCGATAACTTGATG 1410
                 | |||  ||| |
Db        1433 AAGTTAAAGTGAAG 1446

RESULT 9
AAQ74691
```

```
ID   AAQ74691 standard; cDNA; 1783 BP.
XX
AC   AAQ74691;
XX
DT   25-MAR-2003  (updated)
DT   22-JUN-1995  (first entry)
XX
DE   Early Ripening Tomato protein ERTD1 gene.
XX
KW   Early Ripening Tomato; transgenic plant; crop improvement;
KW   fruit ripening; transformation; ss.
XX
OS   Lycopersicon esculentum.
XX
PN   WO9421794-A1.
XX
PD   29-SEP-1994.
XX
PF   22-MAR-1994;   94WO-GB00581.
XX
PR   22-MAR-1993;   93GB-0005859.
PR   22-MAR-1993;   93GB-0005860.
PR   22-MAR-1993;   93GB-0005862.
PR   22-MAR-1993;   93GB-0005865.
PR   22-MAR-1993;   93GB-0005866.
PR   22-MAR-1993;   93GB-0005867.
PR   22-MAR-1993;   93GB-0005868.
PR   22-MAR-1993;   93GB-0005869.
PR   12-JUL-1993;   93GB-0014351.
PR   12-OCT-1993;   93GB-0020988.
XX
PA   (ZENE ) ZENECA LTD.
XX
PI   Abu-Bakar UK,  Barton SL,  Gallego-Veigas PP,  Gray JE;
PI   Grierson D,  Lowe AL,  Picton S,  Whotton LC,  Abubakar UK;
XX
DR   WPI; 1994-317014/39.
XX
PT   DNA constructs encoding fruit-ripening related proteins - useful
PT   for transformation of plants to modify fruit quality
XX
PS   Claim 1; Page 54; 74pp; English.
XX
CC   ERTD1 homologous transcripts are 1.8 kb in size.  The transcript
CC   is found throughout early fruit development of both wild-type and
CC   rin (ripening inhibitor) fruit.  The transcript is detected
CC   throughout the entire ripening period in rin fruit.  The transcript
CC   is not detected in leaves or wounded leaves of either rin or wild-
CC   type tomato.  The sequence has been deposited as NCIMB 40588.  The DNA
CC   may be used to transform plants for production of plants with an altered
CC   level of expression.  Improvements in transgenic plants include
CC   resistance to damage and pests, longer shelf life, improved
CC   flavor/aroma, etc.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 1783 BP; 512 A; 339 C; 423 G; 509 T; 0 other;

  Query Match             48.8%;  Score 736.6;  DB 15;  Length 1783;
  Best Local Similarity   71.8%;  Pred. No. 2e-215;
  Matches  964;  Conservative    0;  Mismatches  379;  Indels    0;  Gaps    0;

Qy         20 TATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCAC 79
              ||   || || || |    | | ||||  ||||| ||||||| | || || |      |
Db         28 TAAGAGATTCAGAAGAGAGCTTGCACTGTACATTTGCATCAAGATATGTACAGGAACCTT 87

Qy         80 TTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCA 139
              | |||| |||||| |||||  |||| || || || || ||||| || ||||||||  | |
Db         88 TACCTAAGTTCAAAATGCCTAAAAAATCCATGCCGAAAGAAGCAGCTTATCAGATTGTAA 147

Qy        140 ACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACAT 199
              || ||| | || || || ||  ||||||| |||| ||||||| |  ||||||||||
Db        148 ACGACGAGCTTATGTTGGATGGTAACCCCAGGTTGAATTTAGCTTCCTTTGTTAGCACAT 207
```

```
     200 GGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACA 259
         |||||||||| ||||| ||||| ||||||||||| ||||| || || |||||||| ||||
     208 GGATGGAGCCCGAGTGCGATAAGCTCATCATGTCATCCATTAATAAAAACTATGTCGACA 267

     260 TGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATC 319
         |||| ||||| || |||||||| |||||||| ||  ||||||| |||||| | |||||||
     268 TGGATGAGTATCCTGTCACCACTGAACTTCAAAATAGATGTGTTAACATGTTAGCACATC 327

     320 TATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCAT 379
         | ||| |||| ||| | |  || |  ||||| || || ||||| || || || || ||||
     328 TTTTCCATGCCCCGGTTGGTGATGATGAGACTGCAGTTGGAGTTGGTACAGTGGGTTCAT 387

     380 CGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAG 439
         | ||||| |||||| | || ||  | || ||||| || ||||||||    ||  | ||||
     388 CAGAGGCAATAATGCTTGCTGGCCTTGCTTTCAAACGCAAATGGCAATCGAAAAGAAAAG 447

     440 CTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTT 499
         | |||||||||||  ||||||| || || || ||||| ||||| ||||| || || || |
     448 CAGAAGGCAAACCTTTCGATAAGCCTAATATAGTCACTGGAGCTAATGTGCAGGTCTGCT 507

     500 GGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGAT 559
         |||| ||||| || ||||| |||||||||||  | |||||| |||||| | |  |||||||
     508 GGGAAAAATTTGCAAGGTATTTTGAGGTTGAGTTGAAGGAGGTGAAACTAAAAGAAGGAT 567

     560 ACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTG 619
         ||||||| |||||||||    |||| || || || || |||||||| || || |||||||
     568 ACTATGTAATGGACCCTGCCAAAGCAGTAGAGATAGTGGATGAGAATACAATATGTGTTG 627

     620 CGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATC 679
         | |  || ||||||| ||||| | ||| || || || ||||| || ||| | ||||| |
     628 CTGCAATCCTTGGTTCTACTCTGACTGGGGAGTTTGAGGATGTGAAGCTCCTAAACGAGC 687

     680 TCTTGGTCGAAAAAGAACAAAGAAACCGGATGGATACACCAATCCACGTGGATGCGGCAA 739
         || |     ||||||||||| ||||||||||||| ||||| || || || ||||| || |
     688 TCCTTACAAAAAAAGAACAAGGAAACCGGATGGAGACACCGATTCATGTCGATGCTGCGA 747

y    740 GTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCT 799
         |||||||||| ||||| || ||  | |  || ||  | |||||||| ||  |  | ||
b    748 GTGGAGGATTTATTGCTCCTTTCCTCTGGCCAGATCTTGAATGGGATTTCCGTTTGCCTC 807

y    800 TGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTT 859
         | ||||| ||||| ||||| || |||||||||||||| ||||| || || ||  | ||||
b    808 TTGTGAAAAGTATAAATGTCAGCGGTCACAAGTATGGCCTTGTATATGCTGGTGTCGGTT 867

y    860 GGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATC 919
         ||||||| ||| | | ||| || || ||||| || |||||| |||| ||||| || || |
b    868 GGGTGATATGGCGGAGCAAGGAAGACTTGCCCGATGAACTCGTCTTTCATATAAACTACC 927

y    920 TTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTG 979
         ||||  |||| || || || ||||||||||| ||||| ||||||||   |||| | ||||
b    928 TTGGGTCTGATCAGCCTACTTTTACTCTCAACTTCTCTAAAGGTTCCTATCAAATAATTG 987

y    980 CTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATT 1039
         | || || || ||  | ||  || | |||   |||||||| |  || || ||| ||||||
b    988 CACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGTTATAAGAACGTCATGAAGAATT 1047

y   1040 GCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCG 1099
         ||  |   ||    |  || ||||  || ||| |    || |  |   ||||| | || |
b   1048 GCTTATCAAACGCAAAAGTACTAACAGAGGGAATCACAAAAATGGGGCGGTTCGATATTG 1107

y   1100 TCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACA 1159
         |||| ||||| | ||| || ||  |||| || || ||  | |  || ||||||    | |
b   1108 TCTCTAAGGATGTGGGTGTTCCTGTTGTAGCATTTTCTCTCAGGGACAGCAGCAAATATA 1167

y   1160 CTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAA 1219
         | |  || ||| | || ||    ||  | || | |||||||||| || || || |||||||
b   1168 CGGTATTTGAAGTATCTGAGCATCTCAGAAGATTTGGATGGATCGTCCCTGCATACACAA 1227

y   1220 TGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGA 1279
         |||| ||  ||||  |||||||  |||| || || || || || ||||| ||||||
b   1228 TGCCACCGGATGCTGAACACATTGCTGTACTGCGGGTTGTCATTAGAGAGGATTTCAGCC 1287

y   1280 GAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGC 1339
```

```
              |  || ||||||||||||||    || || |||||| |  ||   ||| | ||    |
Db     1288 ACAGCCTAGCTGAGAGACTTGTTTCTGACATTGAGAAAATTCTGTCAGAGTTGGACACAC 1347

Qy     1340 TTCCTTCGAGAGTGATTCACAAA 1362
              ||| |  |  ||     ||||
Db     1348 AGCCTCCTCGTTTGCCCACCAAA 1370


RESULT 10
AAL43418
ID   AAL43418 standard; cDNA; 1783 BP.
XX
AC   AAL43418;
XX
DT   25-SEP-2002  (first entry)
XX
DE   Tomato GAD coding sequence.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
XX
OS   Lycopersicon esculentum.
XX
FH   Key             Location/Qualifiers
FT   CDS             6..1514
FT                   /*tag=  a
FT                   /product= "GAD"
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   P-PSDB; AAO15140.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
PT   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT    -
XX
PS   Claim 18; Page 59; 63pp; English.
XX
CC   The present invention relates to a method of producing a transformed
CC   plant that selectively increases production of gamma-aminobutyric acid
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the tomato GAD coding sequence.
XX
SQ   Sequence 1783 BP; 512 A; 339 C; 423 G; 509 T; 0 other;

  Query Match             48.8%;  Score 736.6;  DB 24;  Length 1783;
  Best Local Similarity   71.8%;  Pred. No. 2e-215;
  Matches  964;  Conservative    0;  Mismatches  379;  Indels    0;  Gaps    0;

Qy           20 TATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCACGTTACGTCCGTACTTCAC 79
                ||   || || || |    | | ||||  ||||| |||||| | || || |      |
Db           28 TAAGAGATTCAGAAGAGAGCTTGCACTGTACATTTGCATCAAGATATGTACAGGAACCTT 87

Qy           80 TTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCA 139
                | |||| |||||| |||||  |||| || || || || ||||| || ||||||||  | |
Db           88 TACCTAAGTTCAAAATGCCTAAAAAATCCATGCCGAAAGAAGCAGCTTATCAGATTGTAA 147
```

```
y        140 ACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACAT 199
             |||||||||| ||| | || || || ||  ||||||| ||||| ||||||||| |  ||||
b        148 ACGACGAGCTTATGTTGGATGGTAACCCCAGGTTGAATTTAGCTTCCTTTGTTAGCACAT 207

y        200 GGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACA 259
             |||||||||| ||||| ||||| ||||||||||| ||||| || || |||||||| ||||
b        208 GGATGGAGCCCGAGTGCGATAAGCTCATCATGTCATCCATTAATAAAAACTATGTCGACA 267

y        260 TGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATC 319
             |||| ||||| || |||||||| |||||||| ||  ||||||| |||||| | |||||||
b        268 TGGATGAGTATCCTGTCACCACTGAACTTCAAAATAGATGTGTTAACATGTTAGCACATC 327

y        320 TATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCAT 379
             | ||| |||| ||| | |  || |  ||||| || || ||||| || || || || ||||
b        328 TTTTCCATGCCCCGGTTGGTGATGATGAGACTGCAGTTGGAGTTGGTACAGTGGGTTCAT 387

y        380 CGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAG 439
             | ||||| |||||| | || ||  | || ||||| || ||||||||    ||  | ||||
b        388 CAGAGGCAATAATGCTTGCTGGCCTTGCTTTCAAACGCAAATGGCAATCGAAAAGAAAAG 447

y        440 CTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAATGTTCAAGTGTGTT 499
             | |||||||||||  ||||||| || || || ||||| ||||| ||||| || || || |
b        448 CAGAAGGCAAACCTTTCGATAAGCCTAATATAGTCACTGGAGCTAATGTGCAGGTCTGCT 507

y        500 GGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGAT 559
             |||| ||||| || ||||| |||||||||||  | ||||| |||||| | |  |||||||
b        508 GGGAAAAATTTGCAAGGTATTTTGAGGTTGAGTTGAAGGAGGTGAAACTAAAAGAAGGAT 567

y        560 ACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGATGAGAACACCATTTGTGTTG 619
             ||||||| |||||||||    |||| || || || || |||||||| || || |||||||
b        568 ACTATGTAATGGACCCTGCCAAAGCAGTAGAGATAGTGGATGAGAATACAATATGTGTTG 627

y        620 CGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGATGTTAAACTCTTGAACGATC 679
             | |  || |||||||| ||||| | ||| || || || ||||| || ||| | ||||| |
b        628 CTGCAATCCTTGGTTCTACTCTGACTGGGGAGTTTGAGGATGTGAAGCTCCTAAACGAGC 687

y        680 TCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCAATCCACGTGGATGCGGCAA 739
             || |     |||||||||| |||||||||||||| ||||| || || || ||||| || |
b        688 TCCTTACAAAAAAGAACAAGGAAACCGGATGGGAGACACCGATTCATGTCGATGCTGCGA 747

y        740 GTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCCCT 799
             |||||||||| ||||| || ||  | |  || ||  | |||||||| ||  |  | ||
b        748 GTGGAGGATTTATTGCTCCTTTCCTCTGGCCAGATCTTGAATGGGATTTCCGTTTGCCTC 807

y        800 TGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGGTT 859
             | ||||| ||||| ||||| || |||||||||||||| ||||| || || ||  | ||||
b        808 TTGTGAAAAGTATAAATGTCAGCGGTCACAAGTATGGCCTTGTATATGCTGGTGTCGGTT 867

y        860 GGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTATC 919
             ||||||| ||| | | ||| || || ||||| || |||||| |||| ||||| || || |
b        868 GGGTGATATGGCGGAGCAAGGAAGACTTGCCCGATGAACTCGTCTTTCATATAAACTACC 927

y        920 TTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCAAGTCAAGTCATTG 979
             ||||  |||| || || || ||||||||||| ||||| ||||||||   |||| | ||||
b        928 TTGGGTCTGATCAGCCTACTTTTACTCTCAACTTCTCTAAAGGTTCCTATCAAATAATTG 987

y        980 CTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTACAGAAATGTGATGGAGAATT 1039
             | || || || ||  | ||  || | |||   |||||||| |  || || ||| ||||||
b        988 CACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGTTATAAGAACGTCATGAAGAATT 1047

y       1040 GCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAGACAGAAAGGTTCAACATCG 1099
             ||  |   ||    |  || ||||  || ||| |    || |  |   ||||| | || |
b       1048 GCTTATCAAACGCAAAAGTACTAACAGAGGGAATCACAAAAATGGGGCGGTTCGATATTG 1107

y       1100 TCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACA 1159
             |||| ||||| | ||| || ||  |||| || || ||  | |  || ||||||    | |
b       1108 TCTCTAAGGATGTGGGTGTTCCTGTTGTAGCATTTTCTCTCAGGGACAGCAGCAAATATA 1167

y       1160 CTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAA 1219
             | |  || ||| | || ||    ||  | || | ||||||||| || || || |||||||
b       1168 CGGTATTTGAAGTATCTGAGCATCTCAGAAGATTTGGATGGATCGTCCCTGCATACACAA 1227
```

```
Qy      1220 TGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGA 1279
             |||| ||  ||||  |||||||  |||| || || || || || ||||| ||||||
Db      1228 TGCCACCGGATGCTGAACACATTGCTGTACTGCGGGTTGTCATTAGAGAGGATTTCAGCC 1287

Qy      1280 GAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGAGC 1339
               |  || ||||||||||||||    || || |||||| |  ||   ||| | ||    |
Db      1288 ACAGCCTAGCTGAGAGACTTGTTTCTGACATTGAGAAAATTCTGTCAGAGTTGGACACAC 1347

Qy      1340 TTCCTTCGAGAGTGATTCACAAA 1362
               ||| |  |  ||     ||||
Db      1348 AGCCTCCTCGTTTGCCCACCAAA 1370


RESULT 11
AAI70004
ID   AAI70004 standard; DNA; 1180 BP.
XX
AC   AAI70004;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Tomato glutamate decarboxylase clone T-gad-19 DNA sequence.
XX
KW   Tomato; glutamate decarboxylase; GAD; transgenic plant; T-gad-19; ds.
XX
OS   Lycopersicon esculentum.
XX
PN   JP2001238555-A.
XX
PD   04-SEP-2001.
XX
PF   13-DEC-2000; 2000JP-0404172.
XX
PR   13-DEC-1999;   99JP-0352552.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 2001-610050/70.
XX
PT   Creation of a transgenic plant of improved amino acid composition,
PT   involves transformation with antisense sequences -
XX
PS   Example 1; Page 13-14; 18pp; Japanese.
XX
CC   The present invention relates to a method for creating a transgenic plant
CC   which accumulates a larger amount of at least one amino acid of edible
CC   part than a natural plant of the same species cultured under a same
CC   condition. The method involves transforming a plant with a genetic
CC   structure containing an antisense sequence of a gene sequence encoding
CC   glutamate decarboxylase (GAD) and a controlling sequence which can
CC   express the antisense sequence. The resultant transgenic plant has high
CC   added value. The present sequence is the DNA sequence for clone T-gad-19
CC   of the tomato GAD gene, which was used to illustrate the present
CC   invention.
XX
SQ   Sequence 1180 BP; 372 A; 188 C; 276 G; 344 T; 0 other;

  Query Match             47.1%;  Score 711.4;  DB 22;  Length 1180;
  Best Local Similarity   75.7%;  Pred. No. 9.1e-208;
  Matches  894;  Conservative    0;  Mismatches  286;  Indels    1;  Gaps    1;

Qy       112 CCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGG 171
             || |||||||| || ||||||||  | || || || || ||| | || || |||||| ||
Db         1 CCAAAGGAAGCAGCATATCAGATTCTAAATGATGAACTTATGTTAGATGGAAATCCAAGG 60

Qy       172 TTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATG 231
             ||||| || || || |||||||| ||||||||||| || || ||||| ||| | || |||
Db        61 TTGAATTTGGCATCTTTTGTGACAACATGGATGGAACCAGAATGTGACAAATTGATGATG 120

Qy       232 TCCTCCATCAACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAG 291
                ||||| ||||| || |||||||||||||| || || || |||||||| || ||||||
Db       121 GATTCCATTAACAAAAATTATGTTGACATGGATGAATATCCTGTCACCACTGAGCTTCAG 180
```

```
y   292 AACCGATGTGTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACC 351
        || || ||||| |||||||| || ||| |||| ||||||||  | ||||| |  || ||
b   181 AATCGGTGTGTAAACATGATAGCGCATTTATTTAATGCACCACTTGAAGATGGAGAAACT 240

y   352 GCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTC 411
        || || ||||| ||||| ||||| || || || ||||| ||| | || || ||||||||
b   241 GCAGTTGGAGTTGGAACAGTTGGTTCTTCAGAAGCCATTATGCTTGCTGGATTGGCCTTT 300

y   412 AAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATT 471
        ||| | |||||||| |||||    |||||  |||| || |||   ||||| |||||||||
b   301 AAGAGAAAATGGCAAAACAAAATGAAAGCCCAAGGAAAGCCCTATGATAAGCCCAACATT 360

y   472 GTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAA 531
        || || || || ||||| || ||||||||||| ||||| || ||||| ||||| |||||
b   361 GTTACTGGTGCTAATGTCCAGGTGTGTTGGGAAAAATTTGCAAGGTATTTTGAAGTTGAG 420

y   532 CTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGAT 591
        || || |||||||| |||| ||| |||||||||||||||  |||| |  ||||||| ||
b   421 CTAAAAGAAGTGAAGTTGACTGATGGATACTATGTGATGAGCCCTGAGAAAGCTGTGGAA 480

y   592 ATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAA 651
        ||||||||||||||||| |||||||| || |  ||| | ||||| ||||| |||||  |
b   481 ATGGTTGATGAGAACACAATTTGTGTAGCTGCTATTTTGGGTTCAACTCTGAATGGGGAG 540

y   652 TTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGG 711
        || || || || ||    ||||| || ||| |  | | ||||||||||| ||||| || |||
b   541 TTTGAGGAAGTGAAGAAATTGAATGACCTCCT-ATTGAAAAGAACAAGGAAACAGGGTGG 599

y   712 GATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCG 771
        || || ||||| || |||||||| || ||||| ||||| |||||||| ||| | |||||
b   600 GACACTCCAATTCATGTGGATGCAGCTAGTGGTGGATTTATTGCACCATTTATATATCCA 659

y   772 GAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAG 831
        ||| | ||||||||||||||| | || || || || ||||| ||||| || |||||||||
b   660 GAACTTGAATGGGACTTTAGATTGCCATTAGTTAAAAGTATAAATGTCAGCGGTCACAAG 719

y   832 TATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCT 891
        ||||| ||||| || || ||  | |||||||| || ||| | | ||| || || |||||
b   720 TATGGCCTTGTATATGCTGGTGTCGGTTGGGTCATATGGCGGAGCAAGGAAGACTTGCCC 779

y   892 GAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAAT 951
        || ||||| || || ||||| ||||| |||||  |||| ||||| || || || ||||||
b   780 GATGAACTTATTTTTCATATTAATTACCTTGGGTCTGATCAACCTACTTTCACCCTCAAT 839

y   952 TTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCAC 1011
        ||||| ||||| ||||||||||| ||||||||||| || ||||| || || || ||  |
b   840 TTCTCTAAAGGATCAAGTCAAGTAATTGCTCAATATTATCAACTAATTCGTTTTGGTTAT 899

y  1012 GAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGA 1071
        |||||||| |  || || ||| ||||||||  |   ||    |  || ||||  || |||
b   900 GAGGGTTATAAGAACGTCATGAAGAATTGCTTATCAAACGCAAGGGTACTAACAGAGGGA 959

y  1072 CTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCT 1131
         |    || |  | ||| ||| | || ||||| ||||| |  || || ||  | || ||
b   960 ATCACAAAAATGGGAAGATTCGAGATTGTCTCCAAGGATGTTGGTGTTCCCTTGGTTGCA 1019

y  1132 TTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGG 1191
        || ||  | ||||| |  |||   ||   |||||| ||| | || || |   |  | ||
b  1020 TTTTCTCTTAAAGACAATAGCAAACATGATGAGTTTGAAGTATCTGAAACTTTAAGGAGA 1079

y  1192 TATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTT 1251
        | ||||||||| || || || ||||| ||||| ||| |||| ||||||||||| ||||||
b  1080 TTTGGATGGATTGTTCCAGCATACACTATGCCACCAGATGCTCAACACATCACAGTTCTT 1139

y  1252 CGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCTGA 1292
         | || ||||| ||||||||||||||  | | |||||||| ||
b  1140 AGAGTTGTTATTAGAGAAGATTTCTCCCGTACACTCGCCGA 1180

ESULT 12
BX62354
D   ABX62354 standard; DNA; 416 BP.
X
```

```
AC   ABX62354;
XX
DT   25-FEB-2003  (first entry)
XX
DE   Arabidopsis thaliana expressed sequence related polynucleotide #469.
XX
KW   Transgenic plant; plant; genetically modified cell;
KW   environmental stress; ribozyme creation; disease resistance;
KW   stress tolerance; fungicide screening; insecticide screening;
KW   gene; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   US2002040490-A1.
XX
PD   04-APR-2002.
XX
PF   26-JAN-2001; 2001US-0770423.
XX
PR   27-JAN-2000; 2000US-178512P.
XX
PA   (GORL/) GORLACH J.
PA   (ANYY/) AN Y.
PA   (HAMI/) HAMILTON C M.
PA   (PRIC/) PRICE J L.
PA   (RAIN/) RAINES T M.
PA   (YUYY/) YU Y.
PA   (RAME/) RAMEAKA J G.
PA   (PAGE/) PAGE A.
PA   (MATH/) MATHEW A V.
PA   (LEDF/) LEDFORD B L.
PA   (WOES/) WOESSNER J P.
PA   (HAAS/) HAAS W D.
PA   (GARC/) GARCIA C A.
PA   (KRIC/) KRICKER M.
PA   (SLAT/) SLATER T.
PA   (DAVI/) DAVIS K R.
PA   (ALLE/) ALLEN K.
PA   (HOFF/) HOFFMAN N.
PA   (HURB/) HURBAN P.
XX
PI   Gorlach J,  An Y,  Hamilton CM,  Price JL,  Raines TM,  Yu Y;
PI   Rameaka JG,  Page A,  Mathew AV,  Ledford BL,  Woessner JP,  Haas WD;
PI   Garcia CA,  Kricker M,  Slater T,  Davis KR,  Allen K,  Hoffman N;
PI   Hurban P;
XX
DR   WPI; 2003-110411/10.
XX
PT   Novel Arabidopsis thaliana nucleic acid useful for identifying
PT   homologous or related genes, and to create genetically modified and
PT   transgenic organisms, such as plant cells and plants  -
XX
PS   Claim 1; SEQ ID NO 469; 43pp; English.
XX
CC   The invention describes an Arabidopsis thaliana nucleic acid (I). The
CC   polypeptide (II) encoded by (I), transgenic plant (III) or geneticallly
CC   modified cell (IV) are useful for screening a candidate agent for its
CC   biological effect, by combining the candidate agent with (II), (III) or
CC   (IV), and determining the effect of the candidate agent on (II), (III) or
CC   (IV). (I) is useful for identifying homologous or related genes, for
CC   producing compositions that modulate the expression or function of its
CC   encoded protein, for mapping functional regions of the protein, in
CC   diagnosis, for studying associated physiological pathways, for genetic
CC   manipulation of cells, preferably plant cells, in screening assays of
CC   various plant strains to determine the strains that are capable of
CC   withstanding a particular disease or environmental stress, for enhancing
CC   or inhibiting production of biosynthetic product in a plant, for
CC   producing polypeptides, as probes for the detection of mRNA in biological
CC   samples, to generate additional copies of (I), to generate ribozymes or
CC   oligonucleotides, as single stranded DNA probes or as triple-strand
CC   forming oligonucleotides, and to create genetically modified and
CC   transgenic organisms, such as plant cells and plants. (II) or (III) is
CC   useful for introducing or improving disease resistance and stress
```

```
   tolerance in plants, screening biological active agents, e.g.,
   fungicides, insecticides, etc., and for elucidating biochemical pathways.
   (III) is useful as crops for their enhanced diseased resistance, enhanced
   traits of interest, for screening programs, as crops which exhibit
   enhanced tolerance to environmental stress, or to produce a factor. This
   sequence represents a nucleic acid that may correspond to naturally
   occurring Arabidopsis thaliana expressed sequences.
   Note: The sequence data for this patent did not form part of the printed
   specification, but was obtained in electronic format directly from the
   US patent office at
   seqdata.uspto.gov/sequence.html?DocID=999909770423.
X
Q  Sequence 416 BP; 102 A; 114 C; 102 G; 98 T; 0 other;

 Query Match            26.6%;  Score 402;  DB 25;  Length 416;
 Best Local Similarity  100.0%;  Pred. No. 6.7e-113;
 Matches  402;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         15 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 74

y         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         75 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 134

y        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        135 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 194

y        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        195 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 254

y        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        255 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 314

y        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        315 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 374

y        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGT 402
             ||||||||||||||||||||||||||||||||||||||||||
b        375 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGT 416


ESULT 13
AC44889
D   AAC44889 standard; DNA; 1607 BP.
X
C   AAC44889;
X
T   18-OCT-2000  (first entry)
X
E   Arabidopsis thaliana DNA fragment SEQ ID NO: 44511.
X
W   Hybridisation assay; genetic mapping; gene expression control;
W   protein identification; signal transduction pathway;
W   metabolic pathway; promoter; termination sequence; ss.
X
S   Arabidopsis thaliana.
X
N   EP1033405-A2.
X
D   06-SEP-2000.
X
F   25-FEB-2000; 2000EP-0301439.
X
R   25-FEB-1999;    99US-0121825.
R   05-MAR-1999;    99US-0123180.
R   09-MAR-1999;    99US-0123548.
R   23-MAR-1999;    99US-0125788.
PR  25-MAR-1999;    99US-0126264.
PR  29-MAR-1999;    99US-0126785.
PR  01-APR-1999;    99US-0127462.
PR  06-APR-1999;    99US-0128234.
PR  08-APR-1999;    99US-0128714.
PR  16-APR-1999;    99US-0129845.
PR  19-APR-1999;    99US-0130077.
PR  21-APR-1999;    99US-0130449.
PR  23-APR-1999;    99US-0130510.
PR  23-APR-1999;    99US-0130891.
PR  28-APR-1999;    99US-0131449.
PR  30-APR-1999;    99US-0132048.
PR  30-APR-1999;    99US-0132407.
PR  04-MAY-1999;    99US-0132484.
PR  05-MAY-1999;    99US-0132485.
PR  06-MAY-1999;    99US-0132486.
PR  06-MAY-1999;    99US-0132487.
PR  07-MAY-1999;    99US-0132863.
PR  11-MAY-1999;    99US-0134256.
PR  14-MAY-1999;    99US-0134218.
PR  14-MAY-1999;    99US-0134219.
PR  14-MAY-1999;    99US-0134221.
PR  14-MAY-1999;    99US-0134370.
PR  18-MAY-1999;    99US-0134768.
PR  19-MAY-1999;    99US-0134941.
PR  20-MAY-1999;    99US-0135124.
PR  21-MAY-1999;    99US-0135353.
PR  24-MAY-1999;    99US-0135629.
PR  25-MAY-1999;    99US-0136021.
PR  27-MAY-1999;    99US-0136392.
PR  28-MAY-1999;    99US-0136782.
PR  01-JUN-1999;    99US-0137222.
PR  03-JUN-1999;    99US-0137528.
PR  04-JUN-1999;    99US-0137502.
PR  07-JUN-1999;    99US-0137724.
PR  08-JUN-1999;    99US-0138094.
PR  10-JUN-1999;    99US-0138540.
PR  10-JUN-1999;    99US-0138847.
PR  14-JUN-1999;    99US-0139119.
PR  16-JUN-1999;    99US-0139452.
PR  16-JUN-1999;    99US-0139453.
PR  17-JUN-1999;    99US-0139492.
PR  18-JUN-1999;    99US-0139454.
PR  18-JUN-1999;    99US-0139455.
PR  18-JUN-1999;    99US-0139456.
PR  18-JUN-1999;    99US-0139457.
PR  18-JUN-1999;    99US-0139458.
PR  18-JUN-1999;    99US-0139459.
PR  18-JUN-1999;    99US-0139460.
PR  18-JUN-1999;    99US-0139461.
PR  18-JUN-1999;    99US-0139462.
PR  18-JUN-1999;    99US-0139463.
PR  18-JUN-1999;    99US-0139750.
PR  18-JUN-1999;    99US-0139763.
PR  21-JUN-1999;    99US-0139817.
PR  22-JUN-1999;    99US-0139899.
PR  23-JUN-1999;    99US-0140353.
PR  23-JUN-1999;    99US-0140354.
PR  24-JUN-1999;    99US-0140695.
PR  28-JUN-1999;    99US-0140823.
PR  29-JUN-1999;    99US-0140991.
PR  30-JUN-1999;    99US-0141287.
PR  01-JUL-1999;    99US-0141842.
PR  01-JUL-1999;    99US-0142154.
PR  02-JUL-1999;    99US-0142055.
PR  06-JUL-1999;    99US-0142390.
PR  08-JUL-1999;    99US-0142803.
PR  09-JUL-1999;    99US-0142920.
PR  12-JUL-1999;    99US-0142977.
PR  13-JUL-1999;    99US-0143542.
PR  14-JUL-1999;    99US-0143624.
PR  15-JUL-1999;    99US-0144005.
PR  16-JUL-1999;    99US-0144085.
```

```
    16-JUL-1999;   99US-0144086.
    19-JUL-1999;   99US-0144325.
    19-JUL-1999;   99US-0144331.
    19-JUL-1999;   99US-0144332.
    19-JUL-1999;   99US-0144333.
    19-JUL-1999;   99US-0144334.
    19-JUL-1999;   99US-0144335.
    20-JUL-1999;   99US-0144352.
    20-JUL-1999;   99US-0144632.
    20-JUL-1999;   99US-0144884.
    21-JUL-1999;   99US-0144814.
    21-JUL-1999;   99US-0145086.
    21-JUL-1999;   99US-0145088.
    22-JUL-1999;   99US-0145085.
    22-JUL-1999;   99US-0145087.
    22-JUL-1999;   99US-0145089.
    22-JUL-1999;   99US-0145192.
    23-JUL-1999;   99US-0145145.
    23-JUL-1999;   99US-0145218.
    23-JUL-1999;   99US-0145224.
    26-JUL-1999;   99US-0145276.
    27-JUL-1999;   99US-0145913.
    27-JUL-1999;   99US-0145918.
    27-JUL-1999;   99US-0145919.
    28-JUL-1999;   99US-0145951.
    02-AUG-1999;   99US-0146386.
    02-AUG-1999;   99US-0146388.
    02-AUG-1999;   99US-0146389.
R   03-AUG-1999;   99US-0147038.
R   04-AUG-1999;   99US-0147204.
R   04-AUG-1999;   99US-0147302.
R   05-AUG-1999;   99US-0147192.
R   05-AUG-1999;   99US-0147260.
R   06-AUG-1999;   99US-0147303.
R   06-AUG-1999;   99US-0147416.
R   09-AUG-1999;   99US-0147493.
R   09-AUG-1999;   99US-0147935.
R   10-AUG-1999;   99US-0148171.
R   11-AUG-1999;   99US-0148319.
R   12-AUG-1999;   99US-0148341.
R   13-AUG-1999;   99US-0148565.
R   13-AUG-1999;   99US-0148684.
R   16-AUG-1999;   99US-0149368.
R   17-AUG-1999;   99US-0149175.
R   18-AUG-1999;   99US-0149426.
R   20-AUG-1999;   99US-0149722.
R   20-AUG-1999;   99US-0149723.
R   20-AUG-1999;   99US-0149929.
R   23-AUG-1999;   99US-0149902.
R   23-AUG-1999;   99US-0149930.
R   25-AUG-1999;   99US-0150566.
R   26-AUG-1999;   99US-0150884.
R   27-AUG-1999;   99US-0151065.
R   27-AUG-1999;   99US-0151066.
R   27-AUG-1999;   99US-0151080.
R   30-AUG-1999;   99US-0151303.
R   31-AUG-1999;   99US-0151438.
R   01-SEP-1999;   99US-0151930.
R   07-SEP-1999;   99US-0152363.
R   10-SEP-1999;   99US-0153070.
R   13-SEP-1999;   99US-0153758.
R   15-SEP-1999;   99US-0154018.
R   16-SEP-1999;   99US-0154039.
R   20-SEP-1999;   99US-0154779.
R   22-SEP-1999;   99US-0155139.
R   23-SEP-1999;   99US-0155486.
R   24-SEP-1999;   99US-0155659.
R   28-SEP-1999;   99US-0156458.
R   29-SEP-1999;   99US-0156596.
R   04-OCT-1999;   99US-0157117.
R   05-OCT-1999;   99US-0157753.
R   06-OCT-1999;   99US-0157865.
R   07-OCT-1999;   99US-0158029.
```

```
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             25.8%;  Score 388.6;  DB 21;  Length 1607;
  Best Local Similarity   64.9%;  Pred. No. 1.9e-108;
  Matches  793;  Conservative    0;  Mismatches  284;  Indels  144;  Gaps    9;

Qy          292 AACCGATGTGTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACC 351
                ||||||||||| ||||| || || |  || |||||||| ||  | || ||  | ||||| 
Db          361 AACCGATGTGTAAACATTATAGCTCGACTGTTCAATGCGCCACTCGAGGAATCTGAGACG 420

Qy          352 GCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTC 411
                || || |||||||| || ||||| || || || ||||| ||||| ||||| |||||||||
Db          421 GCGGTGGGAGTAGGGACAGTTGGTTCTTCAGAAGCCATCATGTTAGCCGGATTGGCCTTC 480

Qy          412 AAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATT 471
                ||  | |||||||||||||| ||||| |||||| || ||||||   || ||||||||||||
Db          481 AAAAGAAAATGGCAGAACAAACGCAAGGCTGAGGGTAAACCCTATGACAAACCCAACATT 540

Qy          472 GTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAA 531
                     | |        |  |  |   |||||||||||||||| ||||||| ||||| || 
Db          541 TGTCACCGGATGCCAATCAAGTTTGCTTGGGAGAAATTCGCTCGGTACTTCGAGGTGGAG 600

Qy          532 CTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGAT 591
                |   | | ||| ||  | |||||||| ||||| ||||||||| ||  |  |||| |  || 
Db          601 C---AAGGAGTAAACCTAAGTGAAGGTTACTACGTGATGGATCCAGACAAAGCAGCAGAA 657

Qy          592 ATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAA 651
                ||||| || |||||||| || ||||| ||    ||  | || ||||| ||  || | |   
Db          658 ATGGTAGACGAGAACACAATCTGTGTCGCA---ATATTGGGATCCACACTCAACGCA--- 711

Qy          652 TTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGG 711

Db          712 ------------------------------------------------------------ 711

Qy          712 GATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCG 771
                                               | ||  ||||| || ||||||| | ||||| 
Db          712 ------------------------------AGTGGGGTCATAGCTCCGTTTATCTATCCT 741

Qy          772 GAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAG 831
                ||||| ||||||||||||||| ||||| ||||| ||||| ||||| |||||||||||||||
Db          742 GAATTAGAATGGGACTTTAGGCTTCCTTTGGTTAAGAGCATCAACGTGAGTGGTCACAAG 801

Qy          832 TATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCT 891
                |||||||| || || || || |||||||||||  | |||||      ||||||||| |||
```

```
)     802 TATGGACTAGTCTATGCTGGTATTGGTTGGGTCGTGTGGAGGGCAGCAGAGGATTTACCT 861

'     892 GAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAAT 951
          || || || ||||| ||||| ||||||||||||||||| |||||||| || |||||||||
)     862 GAAGAGCTTATCTTTCATATTAATTATCTTGGTGCTGATCAACCCACTTTCACTCTCAAT 921

(     952 TTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCAC 1011
          |||||| |||| || || ||| | ||||||||||||||||| || ||    |
)     922 TTCTCC-AAGGATCGAGCCAAATTATTGCTCAATACTACCAGCTCATTTCGTCT------ 974

y    1012 GAGGGTTACAGAAATGTGATGGAGAATTGCAGAG-AGAATATGATCGTCCTAAGGGAAGG 1070
                        ||   | |||||||||| || |||| ||| | || || |  |||||
b     975 --------------TGGATTCGAGAATTGCATAGAAGAACATGGTGGTTCTCAAAGAAGG 1020

y    1071 ACTTGAGAAGACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGC 1130
            | ||||| |||||  | |||||||| |||||||||||| | ||||||||| | |||||
)    1021 TATAGAGAAAACAGAGCGTTTCAACATAGTCTCAAAGGACCAAGGAGTGCCAGTCGTCGC 1080

y    1131 TTTCTCCT-TGAAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCA 1189
           |||| || | ||    |  || |  ||||  |||||||| ||||| || ||||| ||
)    1081 CTTCTTCTCTCAAGGACCATAGTTTCCACAACGAGTTCGAGATCTCTGAGATGCTACGTC 1140

y    1190 GGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTC 1249
          | | ||| ||||| || || || ||||| |||||| |  ||||||| |||||||| ||||
b    1141 GTTTTGGCTGGATCGTCCCAGCTTACACTATGCCTGCCGATGCACAGCACATCACGGTTC 1200

y    1250 TTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCTG-AGAGACTTGTGATCGAT 1308
          | ||||| || ||||| ||||||||||| ||||||||||| | ||||||||||    |||
b    1201 TGCGTGTTGTCATCAGGGAAGATTTCTCAAGAACACTCGCGGAAGAGACTTGTTGCTGAT 1260

y    1309 ATAGAGAAAGTGATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCA 1368
          ||   ||| ||| | | |||||| |||    | ||||| |   |
b    1261 ATTTCGAAGGTGCTTCATGAGCTAGATACCTTGCCTTCCAAGAT---------------- 1304

'y   1369 CTTGGACAAGAGAAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGC 1428
               |    ||||||  | |    |  |          |  || | || ||||||||
b    1305 -----ATCTAAGAAGATGGGAATAGAAGGGATCGCGGAAAATGTAAAGGAGAAGAAGATG 1359

'y   1429 GATATCGACAAGCAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAG 1488
          || |  || |  | ||  ||  | ||   ||| |||| |||||||||    || ||||||
b    1360 GAGAAGGAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAGTTTGTGAAGGAGAGGAAG 1419

'y   1489 AAGACGAGTGGTATCTGCTAA 1509
          |||| || |||| | ||||||
b    1420 AAGATGAATGGTGTGTGCTAA 1440

ESULT 14
.AH23436
D   AAH23436 standard; DNA; 1545 BP.
X
.C  AAH23436;
CX
)T  31-JUL-2001  (first entry)
CX
)E  Aspergillus oryzae glutamate decarboxylase gene related DNA SEQ ID NO: 2.
CX
CW  Glutamate decarboxylase; gamma-aminobutyric acid enriched food;
CW  food production; ds.
CX
)S  Aspergillus oryzae.
XX
PN  JP2001054390-A.
XX
PD  27-FEB-2001.
XX
PF  17-AUG-1999;   99JP-0267604.
XX
PR  17-AUG-1999;   99JP-0267604.
XX
PA  (KIKU-) KIKU MASAMUNE SHUZO KK.
XX
DR  WPI; 2001-276401/29.
```

```
XX
PT   Glutamate decarboxylase gene for producing protein for use in preparing
PT   gamma-aminobutyric acid enriched food  -
XX
PS   Disclosure; Page 7; 12pp; Japanese.
XX
CC   The present invention provides a glutamate decarboxylase gene from
CC   Aspergillus oryzae. This can be used to transform an organism which can
CC   then be used in the production of gamma-aminobutyric acid enriched food.
CC   The present sequence is a DNA used in the exemplification of the
CC   invention.
XX
SQ   Sequence 1545 BP; 373 A; 433 C; 411 G; 328 T; 0 other;

  Query Match             25.6%;  Score 386.2;  DB 22;  Length 1545;
  Best Local Similarity   58.5%;  Pred. No. 1e-107;
  Matches  760;  Conservative    0;  Mismatches  513;  Indels   26;  Gaps    4;

Qy     52 TTCGCATCACGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATT 111
          | || | |||| | ||           || |||  |    ||||||||||      |||
Db    112 TACGGAACACGGTTCGCAACAGAACAGCTGCCTCAGACTGAGATGCCGGAGCGGGAGATG 171

Qy    112 CCTAAGGAAGCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGG 171
          || |  || |  || ||||  || ||||| ||||||||||  || || || || |||  |
Db    172 CCCAGAGAGGTCGCCTATCGCATGATCAAGGACGAGCTGAGCCTCGATGGAAACCCAATG 231

Qy    172 TTGAACTTAGCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATG 231
           | ||||| ||    ||||| ||||| |  |||||   |||    || || ||||| |
Db    232 CTCAACTTGGCGAGTTTTGTCACGACCTATATGGAAGATGAAGCCGAAAAGCTCATGACT 291

Qy    232 TCCTCCATCAACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAG 291
             ||| ||| ||||||||   | ||    || ||||| ||     |  | ||  | |||
Db    292 GAATCCTTCAGCAAGAACTTCATCGATTACGAGGAGTATCCTCAGTCAGCGGAGATCCAG 351

Qy    292 AACCGATGTGTGAACATGATTGCACATCTATTCAATGCACCGTTAGA---AGAGGCGGAG 348
          ||||| || || ||||||||||| | ||| || || || ||  |  |    || |  |||
Db    352 AACCGCTGCGTCAACATGATTGCGCGTCTGTTTAACGCGCCAGTTCACAGCGAAGACGAG 411

Qy    349 ACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCC 408
            | |  | ||       ||| | ||||| || ||||| || ||| |||      ||||
Db    412 CACCCGATGGGCACCTCGACCATCGGATCGTCCGAGGCAATCATGCTGGGTACCCTGGCT 471

Qy    409 TTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAAC 468
           | |||||    ||||| ||||||||||| |||||||||||   |  |      || |||
Db    472 ATGAAGCGCCGTTGGCAAAACAAGCGCAAGGCTGAAGGCAAGGACTACTCCCGCCCTAAC 531

Qy    469 ATTGTCACCGGAGCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTT 528
          || ||||       |  ||| || ||||||||||| ||   |||  |||||||||| ||
Db    532 ATCGTCATGAACAGCGCTGTACAGGTGTGTTGGGAAAAGGCCGCACGGTACTTTGACGTC 591

Qy    529 GAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTT 588
          ||        | ||| | |  | ||| |    ||||||||| || || || || || |||
Db    592 GAGGAGCGATATGTGTACTGCACTGAGGACCGCTATGTGATTGATCCCCAGCAGGCGGTT 651

Qy    589 GATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGA 648
          ||  ||||||| ||||||||  ||  || |    |  || || ||  ||||    |  ||
Db    652 GACCTGGTTGACGAGAACACTATCGGTATCTGCGCTATCCTCGGAACCACCTACACCGGT 711

Qy    649 GAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGA 708
          || |  ||||| || |||    | || || ||  || |||||   ||||
Db    712 GAGTATGAAGACGTCAAAGCAATCAATGACCTTCTGATCGAACGCAACAT---------- 761

Qy    709 TGGGATACACCAATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTAT 768
             |||   || ||||| || || || || || || || ||| | || || ||  |  |
Db    762 --CGATGTTCCCATCCATGTTGACGCCGCTAGCGGCGGTTTCGTCGCTCCTTTCATCAAC 819

Qy    769 CCGGAATTGGAATGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCAC 828
          ||  | |||||||||| ||  | || ||   |||       ||||| |||  ||| |||
Db    820 CCTAAGTTGGAATGGGATTTCCGCCTCCCAAAGGTCGTTTCGATCAACGTGTCTGGCCAC 879

Qy    829 AAGTATGGACTTGTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTG 888
          |||||||| || || ||| | ||  | || |||||  ||||| |   |   ||  || |
Db    880 AAGTATGGTCTGGTCTACCCTGGCGTGGGCTGGGTCGTCTGGCGGTCCCCGGAATATCTC 939
```

```
y          889 CCTGAGGAACTCATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTC 948
               ||  |||| || || ||| | ||||| || || || ||||| ||  |||  ||||| ||
b          940 CCCAAGGATCTGATTTTCAACATCAACTACCTAGGAGCTGAACAGGCCAGTTTTACCCTG 999

y          949 AATTTCTCCAAAGGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGC 1008
               || ||||||||| ||| ||  ||| ||||| |  || ||||||||  | |||||| | |||
b         1000 AACTTCTCCAAGGGTGCATCTCAGGTCATCGGCCAGTACTACCAGATGATCCGACTTGGC 1059

y         1009 CACGAGGGTTACAGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAA 1068
                |    || ||| |     | |||  ||     |  |  | |  |  | |  |     ||
b         1060 AAGCGTGGCTACCGCTCCATCATGACCAACATTACTGTTACTGCGGACTTTTTGGCCCAA 1119

y         1069 GGACTTGAGAAGA----------CAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGT 1118
               |  || |||||||          ||  | |   || | |||   | | || |
b         1120 GAGCTGGAGAAGATGGGCTTTATCATCATGAGCCAGCGTCGCGGCCACGGTCTCCCATTG 1179

y         1119 GCCACTTGTCGCTTTCTCCTTGAAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGA 1178
               | | |||   |  | | |  |||  |  ||   | ||   || ||||| |   |  |  |
b         1180 GTCGCTTTCAGACTGCCCGCTGAGCGGGAGGGTCAGTTCGAC-GAGTTTGCCTTGGCACA 1238

y         1179 CATGCTTCGCAGGTATGGATGGATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACA 1238
               |  ||| ||   |  ||| ||||| || || || ||||| ||| | || ||     |  |
b         1239 CCAGCTGCGTGAGCGTGGCTGGATTGTCCCTGCATACACGATGGCGCCCAACAGCAACAA 1298

y         1239 CATCACTGTTCTTCGTGTGGTTATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACT 1298
               | | |   |  | |||||||||| || | || |||||| | |  |  |     ||    ||
b         1299 CCTGAAGCTGATGCGTGTGGTTGTCCGCGAGGATTTCACAAAGAGCCGGTGCGACGCGCT 1358

y         1299 TGTGATCGATATAGAGAAAGTGATGCGTGAGCTCGATGA 1337
                 | |  |||||  ||   |   ||     | | | |||
b         1359 GCTTAGTGATATCAAGTTGGGCCTGAAGACGTTGGGTGA 1397


ESULT 15
AL43413
D   AAL43413 standard; DNA; 2121 BP.
X
C   AAL43413;
X
T   25-SEP-2002  (first entry)
X
E   A thaliana GAD4 coding sequence.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
X
S   Arabidopsis thaliana.
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I   Kinnersley AM,  Turano FJ;
X
R   WPI; 2002-490073/52.
R   P-PSDB; AAO15135.
X
T   Making transformed plants that selectively increase gamma-aminobutyric
T   acid production, by incorporating a DNA construct with a polynucleotide
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T   -
X
S   Claim 18; Page 55; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
```

```
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the A. thaliana GAD4 coding sequence.
XX
SQ   Sequence 2121 BP; 626 A; 412 C; 473 G; 610 T; 0 other;

  Query Match             25.6%;  Score 385.8;  DB 24;  Length 2121;
  Best Local Similarity   65.6%;  Pred. No. 1.6e-107;
  Matches  701;  Conservative    0;  Mismatches  182;  Indels  186;  Gaps    2;

Qy          84 TAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAAGCGGCGTATCAGATCATCAACGA 143
               ||| ||| | ||||| || |||| || ||| || ||||| || || || ||||||||||
Db         270 TAGATTCGAAATGCCTGAGAACTCAATCCCAAAAGAAGCAGCTTACCAAATCATCAACGA 329

Qy         144 CGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTAGCCTCCTTTGTGACGACATGGAT 203
               |||||| ||||| || || || ||| || ||| |||| |||| ||||| |||| ||||||||
Db         330 CGAGCTAATGCTCGATGGTAACCCAAGGCTGAACCTAGCTTCCTTCGTGACCACATGGAT 389

Qy         204 GGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATCAACAAGAACTATGTTGACATGGA 263
               |||||| || ||||| || ||||| |||   |||||||||||||||||| || ||||||||
Db         390 GGAGCCAGAATGTGACAAGCTCATGATGGAGTCCATCAACAAGAACTACGTCGACATGGA 449

Qy         264 CGAGTACCCCGTCACCACCGAACTTCAGAACCGATGTGTGAACATGATTGCACATCTATT 323
               ||||||||| |||||||| || |||||||||||||||||| |||||||| |||| ||| ||
Db         450 CGAGTACCCTGTCACCACTGAGCTTCAGAACCGATGTGTTAACATGATAGCACGTCTCTT 509

Qy         324 CAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGAGTAGGAACCGTTGGATCATCGGA 383
               ||| || ||| | |  || |   ||  | |||| || || || ||||| ||||| |||||
Db         510 CAACGCGCCGCTTGGTGACGGTGAAGCTGCCGTTGGTGTTGGCACCGTCGGATCGTCGGA 569

Qy         384 GGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAATGGCAGAACAAGCGCAAAGCTGA 443
               ||| || |||||||||||||||||| || ||| |  |||||||||| ||||| || ||  |
Db         570 GGCGATTATGTTGGCCGGTTTGGCTTTTAAGAGACAATGGCAGAATAAGCGTAAGGCCCA 629

Qy         444 AGGCAAACCCGTCGATAAACCCAACATTGTCACCGGAGCCAATG---------------- 487
               |||    ||    ||||| ||||| || || ||||| || ||||
Db         630 AGGGCTTCCTTATGATAAGCCCAATATCGTAACCGGTGCTAATGTCCAGGTAAACCAAAA 689

Qy         488 ------------------------------------------------------------ 487

Db         690 CAAAAATTGATGAAATATTAACCAAGACAAAATTGAATTTATCAATCCGGTTAAGTTATA 749

Qy         488 ------------------------TTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTG 523
                                       | || || || |||||||||||||| ||||| || |
Db         750 TGTGTGACTCAATTTCCGGTTCAATACAGGTTTGCTGGGAGAAATTCGCAAGGTATTTCG 809

Qy         524 AGGTTGAACTTAAGGAAGTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAG 583
               | || || ||||||||||||||  | || ||||  || || ||||||||||||  | | |
Db         810 AAGTGGAGCTTAAGGAAGTGAACCTAAGAGAAGACTATTACGTGATGGACCCTGTAAAGG 869

Qy         584 CTGTTGATATGGTTGATGAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTA 643
               | || || ||||| || || ||||| ||||||||| || | ||| || |||||| ||  | |
Db         870 CGGTCGAAATGGTAGACGAAAACACAATTTGTGTCGCTGCCATCCTCGGTTCAACGTTAA 929

Qy         644 ATGGAGAATTCGAAGATGTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAA 703
                 || ||||||||||| ||||| ||| | ||||| ||| | ||||| || |||||  |||
Db         930 CCGGTGAATTCGAAGACGTTAAGCTCCTCAACGACCTCCTTGTCGAGAAAAACAAGCAAA 989

Qy         704 CCG--------------------------------------------------------- 706
               |||
Db         990 CCGGGTAATTAAACCAAACCGAGAAACAAGCTAATATCGATTGTAATCGGTTTGGAGTCC 1049

Qy         707 -----------------------------GATGGGATACACCAATCCACGTGGATGCGGC 737
                                            ||||||| || ||||| |||||||| || ||
Db        1050 GGTTTTAACGTTCTAAAACACAATTTGCAGATGGGACACGCCAATACACGTGGACGCAGC 1109

Qy         738 AAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAATGGGACTTTAGACTTCC 797
                ||||| || || ||||| |||| |||||||||||  |||| ||||| ||  | || ||
Db        1110 GAGTGGTGGGTTTATTGCTCCGTTCTTGTATCCGGAGCTGGAGTGGGATTTCCGGCTACC 1169
```

```
y        798 CTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTTGTGTACGCAGGGATTGG 857
              ||||| |||||||| ||||||||||||||||| || ||  | || ||||| || |||||
b       1170 GTTGGTTAAGAGTATTAATGTGAGTGGTCACAAATACGGTTTGGTTTACGCCGGTATTGG 1229

y        858 TTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTCATCTTCCATATCAATTA 917
             ||||||  | ||||||| ||||   ||||||||||| ||||| |||||||||||||||||
b       1230 TTGGGTTGTATGGAGAACCAAAACCGATTTGCCTGATGAACTTATCTTCCATATCAATTA 1289

y        918 TCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAAGGTTCA 966
             |||||| ||||| ||||| |||||||| ||||| |||||||||||| ||
b       1290 TCTTGGCGCTGATCAACCAACCTTTACACTCAACTTCTCCAAAGGTACA 1338

earch completed: October 22, 2003, 11:12:29
ob time : 444 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

 protein - protein search, using sw model

n on:         October 22, 2003, 13:37:03 ; Search time 38 Seconds
                                           (without alignments)
                                           621.248 Million cell updates/sec

itle:         US-10-006-852-2
erfect score: 2615
equence:      1 MVLSHAVSESDVSVHSTFAS..........DIITGWKKFVADRKKTSGIC 502

oring table:  BLOSUM62
              Gapop 10.0 , Gapext 0.5

earched:      127863 seqs, 47026705 residues

otal number of hits satisfying chosen parameters:      127863

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :       SwissProt_41:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the result being printed,
     and is derived by analysis of the total score distribution.

                                 SUMMARIES
                   %
esult            Query
 No.     Score   Match Length DB  ID                   Description
-----------------------------------------------------------------------------
   1     2607    99.7    502  1  DCE1_ARATH            Q42521 arabidopsis
   2     2250    86.0    500  1  DCE_PETHY             Q07346 petunia hyb
   3     2130    81.5    494  1  DCE2_ARATH            Q42472 arabidopsis
   4     2014    77.0    502  1  DCE_LYCES             P54767 lycopersico
   5     1028    39.3    466  1  DCE_LACLA             Q9cg20 lactococcus
   6     1023    39.1    466  1  DCE_LACLC             O30418 lactococcus
   7     1013    38.7    466  1  DCEA_ECOLI            P80063 escherichia
   8     1013    38.7    466  1  DCEB_ECOLI            P28302 escherichia
   9     1009    38.6    466  1  DCEA_ECO57            P58228 escherichia
  10      963    36.8    467  1  DCEC_LISIN            Q928k4 listeria in
  11    955.5    36.5    462  1  DCEA_LISMO            Q9f5p3 listeria mo
  12    953.5    36.5    467  1  DCEC_LISMO            Q8y4k4 listeria mo
  13      948    36.3    464  1  DCEB_LISIN            Q928r9 listeria in
  14      940    35.9    464  1  DCEB_LISMO            Q9eyw9 listeria mo
  15    887.5    33.9    585  1  DCE_YEAST             Q04792 saccharomyc
  16      225     8.6    396  1  Y050_METJA            Q60358 methanococc
  17      137     5.2    413  1  DCHS_LYCES            P54772 lycopersico
  18    127.5     4.9    500  1  DDC_CATRO             P17770 catharanthu
  19    124.5     4.8    386  1  DCHS_VIBAN            Q56581 vibrio angu
  20    124.5     4.8    510  1  DDC_ACIBA             Q43908 acinetobact
  21      123     4.7    480  1  DDC_HUMAN             P20711 homo sapien
  22      123     4.7   1935  1  MYH7_HUMAN            P12883 homo sapien
  23      123     4.7   1935  1  MYH7_PIG              P79293 sus scrofa
  24    121.5     4.6    511  1  DDC_HAEIN             P71362 haemophilus
  25      119     4.6    510  1  L2AM_DROME            P18486 drosophila
  26    115.5     4.4    656  1  DCHS_RAT              P16453 rattus norv
  27    114.5     4.4    756  1  METE_BUCBP            Q89b24 buchnera ap
  28    113.5     4.3    804  1  YNA4_CAEEL            P45895 caenorhabdi
  29      113     4.3   1934  1  MYH7_MESAU            P13540 mesocricetu
  30      112     4.3    593  1  DCP1_PEA              P51850 pisum sativ
  31      112     4.3   1935  1  MYH7_RAT              P02564 rattus norv
  32    111.5     4.3    377  1  DCHS_KLEPL            P28578 klebsiella
  33    110.5     4.2   1420  1  APX_XENLA             Q01613 xenopus lae
```

```
  34    109.5     4.2    328  1  L2AM_DROSI            P81893 drosophila
  35    108.5     4.1    480  1  DDC_RAT               P14173 rattus norv
  36    108.5     4.1    625  1  HTPG_BACHD            Q9ke51 bacillus ha
  37    107.5     4.1    377  1  DCHS_MORMO            P05034 morganella
  38    107.5     4.1    662  1  DCHS_HUMAN            P19113 homo sapien
  39    105.5     4.0    403  1  DDC_DROLE             O96571 drosophila
  40    105.5     4.0    536  1  CH60_METAC            Q8tgx7 methanosarc
  41      105     4.0    508  1  DDC_MANSE             P48861 manduca sex
  42    104.5     4.0    369  1  DCHS_RHILO            Q98a07 rhizobium l
  43    104.5     4.0    793  1  YEA8_SCHPO            O14073 schizosacch
  44      104     4.0    486  1  DDC_PIG               P80041 sus scrofa
  45      104     4.0    847  1  DCHS_DROME            Q05733 drosophila

                                ALIGNMENTS

RESULT 1
DCE1_ARATH
ID   DCE1_ARATH     STANDARD;      PRT;   502 AA.
AC   Q42521; Q9FFH9;
DT   01-NOV-1997 (Rel. 35, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glutamate decarboxylase 1 (EC 4.1.1.15) (GAD 1).
GN   GAD1 OR GDH1 OR GAD OR AT5G17330 OR MKP11.30 OR MKP11_18.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=95334488; PubMed=7610159;
RA   Arazi T., Baum G., Snedden W.A., Shelp B.J., Fromm H.;
RT   "Molecular and biochemical analysis of calmodulin interactions with
RT   the calmodulin-binding domain of plant glutamate decarboxylase.";
RL   Plant Physiol. 108:551-561(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=97471969; PubMed=9330910;
RA   Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA   Miyajima N., Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT   features of the 1.6 Mb regions covered by twenty physically assigned
RT   P1 clones.";
RL   DNA Res. 4:215-230(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT   SSP consortium (Salk/Stanford/PGEC).";
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES THE PRODUCTION OF GABA. THE CALMODULIN-BINDING
CC       IS CALCIUM-DEPENDENT AND IT IS PROPOSED THAT THIS MAY, DIRECTLY
CC       OR INDIRECTLY, FORM A CALCIUM REGULATED CONTROL OF GABA
CC       BIOSYNTHESIS.
CC   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
CC   -!- COFACTOR: Pyridoxal phosphate.
CC   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
CC       GAD, HDC AND TYRDC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
```

```
R    EMBL; U10034; AAA93132.1; -.
R    EMBL; AB005238; BAB10520.1; -.
R    EMBL; AY094464; AAM19834.1; -.
R    InterPro; IPR002129; Pyridoxal_deC.
R    Pfam; PF00282; pyridoxal_deC; 1.
R    PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
W    Lyase; Decarboxylase; Pyridoxal phosphate; Calmodulin-binding;
W    Multigene family.
T    BINDING     277    277       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
T    DOMAIN      469    502       CALMODULIN-BINDING.
T    CONFLICT    208    208       A -> D (IN REF. 1).
Q    SEQUENCE   502 AA;  57066 MW;  4E8141FF523E0E22 CRC64;

 Query Match             99.7%;  Score 2607;  DB 1;  Length 502;
 Best Local Similarity   99.8%;  Pred. No. 7.8e-175;
 Matches  501;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

y           1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b           1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

y          61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120

y         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180

y         181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
b         181 VKLSEGYYVMDPQQAVDMVDENTICVAAILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240

y         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300

y         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360

y         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420

y         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480

y         481 QRDIITGWKKFVADRKKTSGIC 502
              ||||||||||||||||||||||
b         481 QRDIITGWKKFVADRKKTSGIC 502

ESULT 2
CE_PETHY
D    DCE_PETHY      STANDARD;      PRT;   500 AA.
C    Q07346;
T    01-NOV-1995 (Rel. 32, Created)
T    01-NOV-1995 (Rel. 32, Last sequence update)
T    28-FEB-2003 (Rel. 41, Last annotation update)
E    Glutamate decarboxylase (EC 4.1.1.15) (GAD).
N    GAD.
S    Petunia hybrida (Petunia).
C    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
C    Asteridae; lamiids; Solanales; Solanaceae; Petunia.
X    NCBI_TaxID=4102;
N    [1]
P    SEQUENCE FROM N.A.
C    TISSUE=Petal;
X    MEDLINE=93374956; PubMed=8366104;
A    Baum G., Chen Y., Arazi T., Takatsuji H., Fromm H.;
```

```
RT   "A plant glutamate decarboxylase containing a calmodulin binding
RT   domain. Cloning, sequence, and functional analysis.";
RL   J. Biol. Chem. 268:19610-19617(1993).
CC   -!- FUNCTION: CATALYZES THE PRODUCTION OF GABA. THE CALMODULIN-BINDING
CC       IS CALCIUM-DEPENDENT AND IT IS PROPOSED THAT THIS MAY, DIRECTLY
CC       OR INDIRECTLY, FORM A CALCIUM REGULATED CONTROL OF GABA
CC       BIOSYNTHESIS.
CC   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
CC   -!- COFACTOR: Pyridoxal phosphate.
CC   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
CC       GAD, HDC AND TYRDC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L16797; AAA33709.1; -.
DR   EMBL; L16977; AAA33710.1; -.
DR   PIR; A48767; A48767.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
DR   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
KW   Lyase; Decarboxylase; Pyridoxal phosphate; Calmodulin-binding.
FT   BINDING     277    277       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT   DOMAIN      469    500       CALMODULIN-BINDING.
SQ   SEQUENCE   500 AA;  56726 MW;  72A043CB885AE10D CRC64;

  Query Match             86.0%;  Score 2250;  DB 1;  Length 500;
  Best Local Similarity   84.9%;  Pred. No. 6.9e-150;
  Matches  427;  Conservative   39;  Mismatches   33;  Indels    4;  Gaps    2;

Qy           1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
               ||||  ||:||||:|||||||||||||||||||:||||||||||||||||||||||||||
Db           1 MVLSKTVSQSDVSIHSTFASRYVRTSLPRFKMPDNSIPKEAAYQIINDELMLDGNPRLNL 60

Qy          61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
               |||||||||||||||:| |||||||||||||||||||||||||||||||||||: |||||
Db          61 ASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEDGETAVG 120

Qy         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
               |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
Db         121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy         181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
               ||||||||||||::||:|||||||||| |||||||||||||| |||||||||||||||||
Db         181 VKLSEGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVEKNKETGWDTP 240

Qy         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
               ||||||||||||||:|||||||||||||||||||||||||||||||||:||||:|||:||
Db         241 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRNKDDLPDEL 300

Qy         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
               ||||||||||||||||||||||||||||||||||||:|||:||||||:||  ||||||||
Db         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGYEGYKNVMENCQENASVLREGLEK 360

Qy         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
               | ||||:||: |||||||||||:  | |||||: |||:||||||||||||||||||||||
Db         361 TGRFNIISKEIGVPLVAFSLKDNRQHNEFEISETLRRFGWIVPAYTMPPNAQHITVLRVV 420

Qy         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
               |||||||||||||| |||||: ||| ||:||  |::: :|:: :|   :    ||:| :
Db         421 IREDFSRTLAERLVRDIEKVLHELDTLPARVNAKLAVAEEQAAANGSEVH---KKTDSEV 477

Qy         481 QRDIITGWKKFVAD-RKKTSGIC 502
               | ::|| ||||| : :|||: :|
Db         478 QLEMITAWKKFVEEKKKKTNRVC 500

RESULT 3
```

```
CE2_ARATH
D   DCE2_ARATH     STANDARD;      PRT;   494 AA.
C   Q42472;
T   01-NOV-1997 (Rel. 35, Created)
T   01-NOV-1997 (Rel. 35, Last sequence update)
T   28-FEB-2003 (Rel. 41, Last annotation update)
E   Glutamate decarboxylase 2 (EC 4.1.1.15) (GAD 2).
N   GAD2 OR GDH2 OR AT1G65960 OR F12P19.12.
S   Arabidopsis thaliana (Mouse-ear cress).
C   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
C   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
X   NCBI_TaxID=3702;
N   [1]
P   SEQUENCE FROM N.A.
C   STRAIN=cv. Columbia;
X   MEDLINE=98363649; PubMed=9700069;
A   Zik M., Arazi T., Snedden W.A., Fromm H.;
T   "Two isoforms of glutamate decarboxylase in Arabidopsis are regulated
T   by calcium/calmodulin and differ in organ distribution.";
L   Plant Mol. Biol. 37:967-975(1998).
N   [2]
P   SEQUENCE FROM N.A.
C   STRAIN=cv. Columbia;
X   MEDLINE=97267149; PubMed=9112779;
A   Turano F.J., Thakkar S.S., Fang T., Weisemann J.M.;
T   "Characterization and expression of NAD(H)-dependent glutamate
T   dehydrogenase genes in Arabidopsis.";
L   Plant Physiol. 113:1329-1341(1997).
N   [3]
P   SEQUENCE FROM N.A.
C   STRAIN=cv. Columbia;
X   MEDLINE=21016719; PubMed=11130712;
A   Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
A   White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
A   Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
A   Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
A   Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
A   Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
A   Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
A   Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
A   Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
A   Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
A   Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
A   Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
A   Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
A   Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
A   Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
A   Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
T   "Sequence and analysis of chromosome 1 of the plant Arabidopsis
T   thaliana.";
L   Nature 408:816-820(2000).
C   -!- FUNCTION: CATALYZES THE PRODUCTION OF GABA. THE CALMODULIN-BINDING
C       IS CALCIUM-DEPENDENT AND IT IS PROPOSED THAT THIS MAY, DIRECTLY
C       OR INDIRECTLY, FORM A CALCIUM REGULATED CONTROL OF GABA
C       BIOSYNTHESIS (BY SIMILARITY).
C   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
C   -!- COFACTOR: Pyridoxal phosphate.
C   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
C       GAD, HDC AND TYRDC).
C   --------------------------------------------------------------------------
C   This SWISS-PROT entry is copyright. It is produced through a collaboration
C   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
C   the European Bioinformatics Institute.  There are no  restrictions on  its
C   use  by  non-profit  institutions as long  as its content  is  in  no  way
C   modified and this statement is not removed.  Usage  by  and for commercial
C   entities requires a license agreement (See http://www.isb-sib.ch/announce/
C   or send an email to license@isb-sib.ch).
C   --------------------------------------------------------------------------
R   EMBL; U49937; AAC31617.1; -.
R   EMBL; U46665; AAC33485.1; -.
R   EMBL; AC009513; AAF06056.1; -.
R   PIR; H96683; H96683.
```

```
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
DR   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
KW   Lyase; Decarboxylase; Pyridoxal phosphate; Calmodulin-binding;
KW   Multigene family.
FT   BINDING     276    276       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT   DOMAIN      461    494       CALMODULIN-BINDING (BY SIMILARITY).
SQ   SEQUENCE   494 AA;  56140 MW;  741E83A25DCBC48C CRC64;

  Query Match             81.5%;  Score 2130;  DB 1;  Length 494;
  Best Local Similarity   80.2%;  Pred. No. 1.7e-141;
  Matches  404;  Conservative   46;  Mismatches   42;  Indels   12;  Gaps    3;

Qy        1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
            |||:   : :| || : | ||||||:||:::: ||||||:|||||| |||||||||||||
Db        1 MVLTKTAT-NDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQIIKDELMLDGNPRLNL 59

Qy       61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
            ||||||||||||||||| ||||||||||||||||||||||||:|| ||||||||:|||||
Db       60 ASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNIIARLFNAPLEESETAVG 119

Qy      121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
            ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      120 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGANVQVCWEKFARYFEVELKE 179

Qy      181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
            | |||||||||| :| :|||||||||| |||||||||||||| ||||||:||:||||:||
Db      180 VNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVKKNEETGWNTP 239

Qy      241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
            ||||||||||||||:|||||||||||||||||||||||||||||||||:||  |||||||
Db      240 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRAAEDLPEEL 299

Qy      301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
            ||||||||||||||||||||||||:||||||||||| |||:|||||| |||:||:||:||
Db      300 IFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRLGFEGYKNVMENCIENMVVLKEGIEK 359

Qy      361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
            ||||||||||:|||:||||||| | | |||||:||||:|||||||||| :||||||||||
Db      360 TERFNIVSKDQGVPVVAFSLKDHSFHNEFEISEMLRRFGWIVPAYTMPADAQHITVLRVV 419

Qy      421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
            |||||||||||||| || ||: ||| |||::  |: :         : :   ||:  ::|
Db      420 IREDFSRTLAERLVADISKVLHELDTLPSKISKKMGI---------EGIAENVKEKKMEK 470

Qy      481 Q--RDIITGWKKFVADRKKTSGIC 502
            :   ::| ||:||| :||| :|:|
Db      471 EILMEVIVGWRKFVKERKKMNGVC 494


RESULT 4
DCE_LYCES
ID   DCE_LYCES      STANDARD;      PRT;   502 AA.
AC   P54767;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glutamate decarboxylase (EC 4.1.1.15) (GAD) (ERT D1).
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Ailsa Craig; TISSUE=Pericarp;
RX   MEDLINE=95284363; PubMed=7766895;
RA   Gallego P.P., Whotton L., Picton S., Grierson D., Gray J.E.;
RT   "A role for glutamate decarboxylase during tomato ripening: the
RT   characterisation of a cDNA encoding a putative glutamate
RT   decarboxylase with a calmodulin-binding site.";
RL   Plant Mol. Biol. 27:1143-1151(1995).
CC   -!- FUNCTION: CATALYZES THE PRODUCTION OF GABA. THE CALMODULIN-BINDING
```

```
C       IS CALCIUM-DEPENDENT AND IT IS PROPOSED THAT THIS MAY, DIRECTLY
C       OR INDIRECTLY, FORM A CALCIUM REGULATED CONTROL OF GABA
C       BIOSYNTHESIS (BY SIMILARITY).
C   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
C   -!- COFACTOR: Pyridoxal phosphate.
C   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
C       GAD, HDC AND TYRDC).
C   --------------------------------------------------------------------------
C   This SWISS-PROT entry is copyright. It is produced through a collaboration
C   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
C   the European Bioinformatics Institute.  There are no  restrictions on  its
C   use  by  non-profit  institutions as long  as its content  is  in  no  way
C   modified and this statement is not removed.  Usage  by  and for commercial
C   entities requires a license agreement (See http://www.isb-sib.ch/announce/
C   or send an email to license@isb-sib.ch).
C   --------------------------------------------------------------------------
R   EMBL; X80840; CAA56812.1; -.
R   PIR; S56177; S56177.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
R   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
W   Lyase; Decarboxylase; Pyridoxal phosphate; Calmodulin-binding.
T   BINDING     278    278       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
T   DOMAIN      471    502       CALMODULIN-BINDING (BY SIMILARITY).
Q   SEQUENCE   502 AA;  56785 MW;  1C5F9BD0084272A6 CRC64;

 Query Match             77.0%;  Score 2014;  DB 1;  Length 502;
 Best Local Similarity   76.0%;  Pred. No. 2.2e-133;
 Matches  387;  Conservative   53;  Mismatches   55;  Indels   14;  Gaps    5;

y           1 MVL-SHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLN 59
              ||| : :: :|: |:| |||||||:  ||:||||: |:||||||||:|||||||||||||
b           1 MVLTTTSIRDSEESLHCTFASRYVQEPLPKFKMPKKSMPKEAAYQIVNDELMLDGNPRLN 60

y          60 LASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAV 119
              |||||:||||||||||||||||||||||||||||||||||:||||:||: : ||||
b          61 LASFVSTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMLAHLFHAPVGDDETAV 120

y         120 GVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELK 179
              |||||||||||||||||||||||:|||||||| |||||||||||||||||||||||||||
b         121 GVGTVGSSEAIMLAGLAFKRKWQSKRKAEGKPFDKPNIVTGANVQVCWEKFARYFEVELK 180

y         180 EVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDT 239
              |||| |||||||| :||::||||||||| |||||| ||||||||||:|| :|||||||:|
b         181 EVKLKEGYYVMDPAKAVEIVDENTICVAAILGSTLTGEFEDVKLLNELLTKKNKETGWET 240

y         240 PIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEE 299
              ||||||||||||||||:|:||||||||||||||||||||||||||:||||||:|||||:|
b         241 PIHVDAASGGFIAPFLWPDLEWDFRLPLVKSINVSGHKYGLVYAGVGWVIWRSKEDLPDE 300

y         300 LIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLE 359
              |:|||||||:||||||||||||| |:||||||||||| |||:|||:||  |  || ||:
b         301 LVFHINYLGSDQPTFTLNFSKGSYQIIAQYYQLIRLGFEGYKNVMKNCLSNAKVLTEGIT 360

y         360 KTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRV 419
              |  ||:||||| |||:|||||:||| :| ||:|:|: |||:|||||||||||:|:|| ||||
b         361 KMGRFDIVSKDVGVPVVAFSLRDSSKYTVFEVSEHLRRFGWIVPAYTMPPDAEHIAVLRV 420

y         420 VIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKI--SLGQEKSESNSDNL----MVTV 473
              ||||||| :|||||| ||||:: |||  | |:  |      :|  :   | |    | ||
b         421 VIREDFSHSLAERLVSDIEKILSELDTQPPRLPTKAVRVTAEEVRDDKGDGLHHFHMDTV 480

y         474 KKSDIDKQRDIITGWKKFVADRKKTSGIC 502
                   : |:|||  |:|     |||||:|
b         481 -----ETQKDIIKHWRKIAG--KKTSGVC 502

ESULT 5
CE_LACLA
D   DCE_LACLA      STANDARD;      PRT;   466 AA.
C   Q9CG20; O50645;
T   16-OCT-2001 (Rel. 40, Created)
T   16-OCT-2001 (Rel. 40, Last sequence update)
```

```
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glutamate decarboxylase (EC 4.1.1.15) (GAD).
GN   GADB OR LL1290.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 1-15, AND CHARACTERIZATION.
RC   STRAIN=01-7;
RX   MEDLINE=99337071; PubMed=10411264;
RA   Nomura M., Nakajima I., Fujita Y., Kobayashi M., Kimoto H., Suzuki I.,
RA   Aso H.;
RT   "Lactococcus lactis contains only one glutamate decarboxylase gene.";
RL   Microbiology 145:1375-1380(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IL1403;
RX   MEDLINE=21235186; PubMed=11337471;
RA   Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT   "The complete genome sequence of the lactic acid bacterium Lactococcus
RT   lactis ssp. lactis IL1403.";
RL   Genome Res. 11:731-753(2001).
CC   -!- FUNCTION: CONVERTS INTERNALIZED GLUTAMATE TO GABA AND INCREASES
CC       THE INTERNAL PH. INVOLVED IN GLUTAMATE-DEPENDENT ACID RESISTANCE
CC       (By similarity).
CC   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
CC   -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC   -!- MISCELLANEOUS: The enzyme is maximally active at pH 4.7. The
CC       activity is stable at acidic pH values; there is no activity in
CC       the neutral pH range. At pH 4.1 the enzyme activity iss retained
CC       at temperatures up to 70 degrees Celsius.
CC   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
CC       GAD, HDC AND TYRDC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB010789; BAA24585.1; -.
DR   EMBL; AE006361; AAK05388.1; -.
DR   PIR; B86786; B86786.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
DR   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
KW   Lyase; Decarboxylase; Pyridoxal phosphate; Complete proteome.
FT   BINDING     277    277       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   466 AA;  53926 MW;  BCDC732253E473C2 CRC64;

  Query Match             39.3%;  Score 1028;  DB 1;  Length 466;
  Best Local Similarity   45.0%;  Pred. No. 1.6e-64;
  Matches  195;  Conservative   91;  Mismatches  141;  Indels    6;  Gaps    3;

Qy          18 FASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIM 77
               | |   :  ||::|: : ||    |||:: ||:: :|| |||||:|  |:||||  ||:
Db          18 FGSESEQVDLPKYKLAQQSIEPRVAYQLVQDEMLDEGNARLNLATFCQTYMEPEAVKLMS 77

Qy          78 SSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
                :: || :|  ||| |||::|||||||| |:||   | |  :|  |:||||| || |:|
Db          78 QTLEKNAIDKSEYPRTTEIENRCVNMIADLWNA--SEKEKFMGTSTIGSSEACMLGGMAM 135

Qy         138 KRKWQNKRKAEGKPVD--KPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
               |  |: : :  |  ::  |||:|  :  |||||||  |:::|::|| : : :  ::  :
Db         136 KFSWRKRAEKLGLDINAKKPNLVISSGYQVCWEKFCIYWDIEMREVPMDKEHMSINLDKV 195

Qy         196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
               :| ||| || |  |:| |  | ::|:| |::|: | ||:| :   |||||||||  |||:
Db         196 MDYVDEYTIGVVGIMGITYTGRYDDIKALDNLIEEYNKQTDYKVYIHVDAASGGLYAPFV 255
```

```
        256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
             |||||||||  | ||| ||||||||| |:|||:||:|: ||||||| ::||| : ||
        256 EPELEWDFRLKNVISINTSGHKYGLVYPGVGWVLWRDKKYLPEELIFKVSYLGGELPTMA 315

        316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPL 375
            :|||  :||:| |||  :| | :||: : |   :  : | : :|||  | |::    :|:
        316 INFSHSASQLIGQYYNFVRYGFDGYKAIHERTHKVAMFLAKEIEKTGMFEIMNDGSQLPI 375

        376 VAFSLKDSSCH--TEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERL 433
            | : ||: |      ::::| |   || |||| :| | ::  : |:||| ||   :|
        376 VCYKLKEDSNRGWNLYDLADRLLMKGWQVPAYPLPKNLENEIIQRLVIRADFGMNMAFNY 435

        434 VIDIEKVMRELDE 446
            | |::: :  |::
        436 VQDMQEAIEALNK 448


ESULT 6
CE_LACLC
D   DCE_LACLC      STANDARD;      PRT;   466 AA.
C   O30418;
T   16-OCT-2001 (Rel. 40, Created)
T   16-OCT-2001 (Rel. 40, Last sequence update)
T   28-FEB-2003 (Rel. 41, Last annotation update)
E   Glutamate decarboxylase (EC 4.1.1.15).
N   GADB.
S   Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
C   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
X   NCBI_TaxID=1359;
N   [1]
P   SEQUENCE FROM N.A., AND CHARACTERIZATION.
C   STRAIN=MG1363;
X   MEDLINE=98143417; PubMed=9484886;
A   Sanders J.W., Leenhouts K., Burghoorn J., Brands J.R., Venema G.,
A   Kok J.;
T   "A chloride-inducible acid resistance mechanism in Lactococcus lactis
T   and its regulation.";
L   Mol. Microbiol. 27:299-310(1998).
C   -!- FUNCTION: CONVERTS INTERNALIZED GLUTAMATE TO GABA AND INCREASES
C       THE INTERNAL PH. INVOLVED IN GLUTAMATE-DEPENDENT ACID RESISTANCE.
C   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
C   -!- COFACTOR: Pyridoxal phosphate (By similarity).
C   -!- INDUCTION: EXPRESSION IS HIGHEST AT ONSET OF STATIONARY PHASE IN
C       PRESENCE OF NACL AND GLUTAMATE, AND AT LOW PH. CHLORIDE-DEPENDENT
C       EXPRESSION IS ACTIVATED BY GADR.
C   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
C       GAD, HDC AND TYRDC).
C   --------------------------------------------------------------------------
C   This SWISS-PROT entry is copyright. It is produced through a collaboration
C   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
C   the European Bioinformatics Institute.  There are no  restrictions on  its
C   use  by  non-profit  institutions as long  as its content  is  in  no  way
C   modified and this statement is not removed.  Usage  by  and for commercial
C   entities requires a license agreement (See http://www.isb-sib.ch/announce/
C   or send an email to license@isb-sib.ch).
C   --------------------------------------------------------------------------
R   EMBL; AF005098; AAC46188.1; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
R   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
W   Lyase; Decarboxylase; Pyridoxal phosphate.
T   BINDING      277    277       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
Q   SEQUENCE   466 AA;  53849 MW;  28B94EB3FEEB168D CRC64;

 Query Match             39.1%;  Score 1023;  DB 1;  Length 466;
 Best Local Similarity   45.0%;  Pred. No. 3.6e-64;
 Matches  195;  Conservative   90;  Mismatches  142;  Indels    6;  Gaps    3;

y          18 FASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIM 77
              | |   :  ||::|: : ||    |||:: ||:: :|| |||||:|  |:||||  ||:
b          18 FGSESEQVDLPKYKLAQQSIEPRVAYQLVQDEMLDEGNARLNLATFCQTYMEPEAVKLMS 77

y          78 SSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
```

```
                :: || :|  ||| |||::|||||||| |:||   |     |  |:||||| || |:|
Db           78 QTLEKNAIDKSEYPRTTEIENRCVNMIADLWNA--SEKGKIYGTSTIGSSEACMLGGMAM 135

Qy          138 KRKWQNKRKAEGKPVD--KPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
                |  |: : :  |  ::  |||:|  :  |||||||  |:::|::|| :   :  :: ::
Db          136 KFSWRKRAEKLGLDINAKKPNLVISSGYQVCWEKFCVYWDIEMREVPMDREHMSINLEKV 195

Qy          196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
                :| ||| || |  |:| |  | ::|:| |::|: | ||:| :   |||||||||  |||:
Db          196 MDYVDEYTIGVVGIMGITYTGRYDDIKALDNLIEEYNKQTDYKVYIHVDAASGGLYAPFV 255

Qy          256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
                 |||||||||  | ||| ||||||||| |:|||:||:|: ||||||| ::||| : ||
Db          256 EPELEWDFRLKNVISINTSGHKYGLVYPGVGWVLWRDKKYLPEELIFKVSYLGGELPTMA 315

Qy          316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPL 375
                :|||  :||:| |||  :| | :||: : |   :  : | | :|||  | |::    :|:
Db          316 INFSHSASQLIGQYYNFVRYGFDGYKAIHERTHKVAMYLAEEIEKTGMFEIMNDGAQLPI 375

Qy          376 VAFSLKDSSCH--TEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERL 433
                | : ||::|      ::::| |   || |||| :| | ::  : |:||| ||   :|
Db          376 VCYKLKENSNRGWNLYDLADRLLMKGWQVPAYPLPKNLENEIIQRLVIRADFGMNMAFNY 435

Qy          434 VIDIEKVMRELDE 446
                | |::: :  |::
Db          436 VQDMQEAIDALNK 448


RESULT 7
DCEA_ECOLI
ID   DCEA_ECOLI     STANDARD;      PRT;   466 AA.
AC   P80063;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glutamate decarboxylase alpha (EC 4.1.1.15) (GAD-alpha).
GN   GADA OR GADS OR B3517 OR C4328.
OS   Escherichia coli, and
OS   Escherichia coli O6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 217992;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=92394884; PubMed=1522060;
RA   Smith D.K., Kassam T., Singh B., Elliott J.F.;
RT   "Escherichia coli has two homologous glutamate decarboxylase genes
RT   that map to distinct loci.";
RL   J. Bacteriol. 174:5820-5826(1992).
RN   [2]
RP   SEQUENCE, AND SEQUENCE OF 148-466 FROM N.A.
RC   STRAIN=ATCC 11246;
RX   MEDLINE=92155241; PubMed=1740158;
RA   Maras B., Sweeney G., Barra D., Bossa F., John R.A.;
RT   "The amino acid sequence of glutamate decarboxylase from Escherichia
RT   coli. Evolutionary relationship between mammalian and bacterial
RT   enzymes.";
RL   Eur. J. Biochem. 204:93-98(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=94316500; PubMed=8041620;
RA   Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT   "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT   region from 76.0 to 81.5 minutes.";
RL   Nucleic Acids Res. 22:2576-2586(1994).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX   MEDLINE=22388234; PubMed=12471157;
RA   Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA   Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
```

```
   Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
   Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
   "Extensive mosaic structure revealed by the complete genome sequence
   of uropathogenic Escherichia coli.";
   Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
   [5]
   SEQUENCE OF 1-22.
   MEDLINE=94033862; PubMed=7764225;
   Yoshida T., Yamashino T., Ueguchi C., Mizuno T.;
   "Expression of the Escherichia coli dimorphic glutamic acid
   decarboxylases is regulated by the nucleoid protein H-NS.";
   Biosci. Biotechnol. Biochem. 57:1568-1569(1993).
   [6]
   SEQUENCE OF 382-392.
   STRAIN=K12 / EMG2;
   MEDLINE=97443975; PubMed=9298646;
   Link A.J., Robison K., Church G.M.;
   "Comparing the predicted and observed properties of proteins encoded
   in the genome of Escherichia coli K-12.";
   Electrophoresis 18:1259-1313(1997).
   -!- FUNCTION: CATALYZES THE PRODUCTION OF GABA.
   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
   -!- COFACTOR: Pyridoxal phosphate.
   -!- SUBUNIT: Homohexamer.
   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
       GAD, HDC AND TYRDC).
   -!- DATABASE: NAME=Worthington enzyme manual;
       WWW="http://www.worthington-biochem.com/manual/G/GLDP.html".
   --------------------------------------------------------------------------
   This SWISS-PROT entry is copyright. It is produced through a collaboration
   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
   the European Bioinformatics Institute.  There are no  restrictions on  its
   use  by  non-profit  institutions as long  as its content  is  in  no  way
   modified and this statement is not removed.  Usage  by  and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
   --------------------------------------------------------------------------
R  EMBL; M84024; AAA23833.1; -.
R  EMBL; X63123; CAA44834.1; -.
R  EMBL; U00039; AAB18493.1; -.
R  EMBL; AE000428; AAC76542.1; -.
R  EMBL; AE016768; AAN82764.1; ALT_INIT.
R  PIR; S47737; S24234.
R  ECO2DBASE; D046.5; 6TH EDITION.
R  ECO2DBASE; E046.5; 6TH EDITION.
R  EcoGene; EG50009; gadA.
R  InterPro; IPR002129; Pyridoxal_deC.
R  Pfam; PF00282; pyridoxal_deC; 1.
R  PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
W  Lyase; Decarboxylase; Pyridoxal phosphate; Multigene family;
W  Complete proteome.
T  BINDING     276    276       PYRIDOXAL PHOSPHATE.
T  CONFLICT     64     64       C -> S (IN REF. 2).
T  CONFLICT     73     73       H -> R (IN REF. 2).
T  CONFLICT    153    153       D -> N (IN REF. 2; AA SEQUENCE).
T  CONFLICT    165    165       C -> S (IN REF. 2).
T  CONFLICT    208    208       T -> N (IN REF. 2).
T  CONFLICT    295    295       L -> V (IN REF. 2).
T  CONFLICT    355    355       D -> N (IN REF. 2; AA SEQUENCE).
Q  SEQUENCE   466 AA;  52685 MW;  86F963E710553E22 CRC64;

 Query Match             38.7%;  Score 1013;  DB 1;  Length 466;
 Best Local Similarity   45.0%;  Pred. No. 1.8e-63;
 Matches  197;  Conservative   83;  Mismatches  152;  Indels    6;  Gaps    3;

y           16 STFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKL 75
               | | :: : |     : | : :  : |:||||||| |||| | |||:|  || :    ||
b           16 SRFGAKAISTIAESKRFPLHEMRDDVAFQIINDELYLDGNARQNLATFCQTWDDENVHKL 75

y           76 IMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGL 135
               :  |||||::| :||| :  :  |||||:| |::||  :   |||  |:||||| || |:
b           76 MDLSINKNWIDKEEYPQSAAIDLRCVNMVADLWHAPAPKNGQAVGTNTIGSSEACMLGGM 135
```

```
Qy         136 AFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
               | | :|: : :| ||| ||||:| |  ||:|| |||||::|||:|: :  |   |||::
Db         136 AMKWRWRKRMEAAGKPTDKPNLVCGP-VQICWHKFARYWDVELREIPMRPGQLFMDPKRM 194

Qy         196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
               ::  ||||| |    | |  | :|  : |:| | :   :|| |  :|:|||||||:|||:
Db         195 IEACDENTIGVVPTFGVTYTGNYEFPQPLHDALDKFQADTGIDIDMHIDAASGGFLAPFV 254

Qy         256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
                |:: |||||| ||||: ||||:||   | ||||||::| ||:||:|:::|||    ||
Db         255 APDIVWDFRLPRVKSISASGHKFGLAPLGCGWVIWRDEEALPQELVFNVDYLGGQIGTFA 314

Qy         316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIV---SKDEG 372
               :|||: : |||||||: :||| |||  |     :    | : : |   :  :     |||
Db         315 INFSRPAGQVIAQYYEFLRLGREGYTKVQNASYQVAAYLADEIAKLGPYEFICTGRPDEG 374

Qy         373 VPLVAFSLKDSS--CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLA 430
               :| | | |||     :| :::|: ||  || |||:|:   |  | |:|:: |  |    |
Db         375 IPAVCFKLKDGEDPGYTLYDLSERLRLRGWQVPAFTLGGEATDIVVMRIMCRRGFEMDFA 434

Qy         431 ERLVIDIEKVMRELDELP 448
               | |: | :  :: | : |
Db         435 ELLLEDYKASLKYLSDHP 452

RESULT 8
DCEB_ECOLI
ID   DCEB_ECOLI     STANDARD;      PRT;   466 AA.
AC   P28302; P76873;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glutamate decarboxylase beta (EC 4.1.1.15) (GAD-beta).
GN   GADB OR B1493 OR Z2215 OR ECS2098 OR SF1734.
OS   Escherichia coli,
OS   Escherichia coli O157:H7, and
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 83334, 623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=K12;
RX   MEDLINE=92394884; PubMed=1522060;
RA   Smith D.K., Kassam T., Singh B., Elliott J.F.;
RT   "Escherichia coli has two homologous glutamate decarboxylase genes
RT   that map to distinct loci.";
RL   J. Bacteriol. 174:5820-5826(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=E.coli; STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA   Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA   Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA   Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA   Oshima T., Saito N., Sampei G., Seki Y., Sivasundaram S.,
RA   Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA   Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
RN   [4]
```

```
     SEQUENCE FROM N.A.
     SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
     MEDLINE=21074935; PubMed=11206551;
     Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
     Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
     Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
     Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
     Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
     Welch R.A., Blattner F.R.;
     "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
     Nature 409:529-533(2001).
     [5]
     SEQUENCE FROM N.A.
     SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
     MEDLINE=21156231; PubMed=11258796;
     Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
     Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
     Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
     Kuhara S., Shiba T., Hattori M., Shinagawa H.;
     "Complete genome sequence of enterohemorrhagic Escherichia coli
     O157:H7 and genomic comparison with a laboratory strain K-12.";
     DNA Res. 8:11-22(2001).
     [6]
     SEQUENCE OF 1-318 FROM N.A.
     SPECIES=E.coli; STRAIN=K12;
     Turlin E., Gasser F., Biville F.;
     "Sequence and functional analysis of an Escherichia coli DNA fragment
     able to complement pqqE and pqqF from Methylobacterium organophilum.";
     Submitted (MAY-1993) to the EMBL/GenBank/DDBJ databases.
     [7]
     SEQUENCE OF 1-15.
     SPECIES=E.coli; STRAIN=K12;
     MEDLINE=93204884; PubMed=8455549;
     Yoshida T., Ueguchi C., Yamada H., Mizuno T.;
     "Function of the Escherichia coli nucleoid protein, H-NS: molecular
     analysis of a subset of proteins whose expression is enhanced in a
     hns deletion mutant.";
     Mol. Gen. Genet. 237:113-122(1993).
     [8]
     SEQUENCE FROM N.A.
     SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
     MEDLINE=22272406; PubMed=12384590;
     Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
     Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
     Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
     Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
     Yu J.;
     "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
     through comparison with genomes of Escherichia coli K12 and O157.";
     Nucleic Acids Res. 30:4432-4441(2002).
     -!- FUNCTION: CATALYZES THE PRODUCTION OF GABA.
     -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
     -!- COFACTOR: Pyridoxal phosphate.
     -!- SUBUNIT: Homohexamer.
     -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
         GAD, HDC AND TYRDC).
     --------------------------------------------------------------------------
     This SWISS-PROT entry is copyright. It is produced through a collaboration
     between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
     the European Bioinformatics Institute.  There are no  restrictions on  its
     use  by  non-profit  institutions as long  as its content  is  in  no  way
     modified and this statement is not removed.  Usage  by  and for commercial
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
     or send an email to license@isb-sib.ch).
     --------------------------------------------------------------------------
     EMBL; M84025; AAA23834.1; -.
     EMBL; AE000246; AAC74566.1; -.
     EMBL; D90791; BAA15163.1; -.
     EMBL; D90790; BAA15157.1; -.
     EMBL; AE005356; AAG56275.1; -.
     EMBL; AP002557; BAB35521.1; -.
     EMBL; X71917; CAA50736.1; ALT_SEQ.
     EMBL; AE015194; AAN43309.1; ALT_INIT.
```

```
DR   PIR; B43332; B43332.
DR   PIR; B90891; B90891.
DR   PIR; G85726; G85726.
DR   EcoGene; EG11490; gadB.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
DR   PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
KW   Lyase; Decarboxylase; Pyridoxal phosphate; Multigene family;
KW   Complete proteome.
FT   BINDING     276    276       PYRIDOXAL PHOSPHATE.
SQ   SEQUENCE   466 AA;  52668 MW;  8E653330A3C5B4ED CRC64;

  Query Match             38.7%;  Score 1013;  DB 1;  Length 466;
  Best Local Similarity   45.0%;  Pred. No. 1.8e-63;
  Matches  197;  Conservative   83;  Mismatches  152;  Indels    6;  Gaps    3;

Qy         16 STFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKL 75
              | | :: : |     : | : :  : |:||||||| |||| | |||:|  || :    ||
Db         16 SRFGAKSISTIAESKRFPLHEMRDDVAFQIINDELYLDGNARQNLATFCQTWDDENVHKL 75

Qy         76 IMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGL 135
              :  |||||::| :||| :  :  |||||:| |::||   :  |||  |:||||| || |:
Db         76 MDLSINKNWIDKEEYPQSAAIDLRCVNMVADLWHAPAPKNGQAVGTNTIGSSEACMLGGM 135

Qy        136 AFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
              | | :|: : :| ||| ||||:| |  ||:|| |||||::|||:|: :  |   |||::
Db        136 AMKWRWRKRMEAAGKPTDKPNLVCGP-VQICWHKFARYWDVELREIPMRPGQLFMDPKRM 194

Qy        196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
              ::  ||||| |    | |  | :|  : |:| | :   :|| |  :|:|||||||:|||:
Db        195 IEACDENTIGVVPTFGVTYTGNYEFPQPLHDALDKFQADTGIDIDMHIDAASGGFLAPFV 254

Qy        256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
               |:: |||||| ||||: ||||:||   | ||||||::| ||:||:|:::|||    ||
Db        255 APDIVWDFRLPRVKSISASGHKFGLAPLGCGWVIWRDEEALPQELVFNVDYLGGQIGTFA 314

Qy        316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIV---SKDEG 372
              :|||: : |||||||: :||| |||  |      :   | : : |   :  :     |||
Db        315 INFSRPAGQVIAQYYEFLRLGREGYTKVQNASYQVAAYLADEIAKLGPYEFICTGRPDEG 374

Qy        373 VPLVAFSLKDSS--CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLA 430
              :| | | |||     :| :::|: ||  || |||:|:   |  | |:|:: |  |    |
Db        375 IPAVCFKLKDGEDPGYTLYDLSERLRLRGWQVPAFTLGGEATDIVVMRIMCRRGFEMDFA 434

Qy        431 ERLVIDIEKVMRELDELP 448
              | |: | :  :: | : |
Db        435 ELLLEDYKASLKYLSDHP 452


RESULT 9
DCEA_ECO57
ID   DCEA_ECO57     STANDARD;      PRT;   466 AA.
AC   P58228;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glutamate decarboxylase alpha (EC 4.1.1.15) (GAD-alpha).
GN   GADA OR GADS OR Z4930 OR ECS4397.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
```

```
     Nature 409:529-533(2001).
     [2]
     SEQUENCE FROM N.A.
     STRAIN=O157:H7 / RIMD 0509952;
     MEDLINE=21156231; PubMed=11258796;
     Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
     Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
     Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
     Kuhara S., Shiba T., Hattori M., Shinagawa H.;
     "Complete genome sequence of enterohemorrhagic Escherichia coli
     O157:H7 and genomic comparison with a laboratory strain K-12.";
     DNA Res. 8:11-22(2001).
     -!- FUNCTION: CATALYZES THE PRODUCTION OF GABA (BY SIMILARITY).
     -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
     -!- COFACTOR: Pyridoxal phosphate (By similarity).
     -!- SUBUNIT: Homohexamer (By similarity).
     -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
         GAD, HDC AND TYRDC).
     --------------------------------------------------------------------------
     This SWISS-PROT entry is copyright. It is produced through a collaboration
     between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
     the European Bioinformatics Institute.  There are no  restrictions on  its
     use  by  non-profit  institutions as long  as its content  is  in  no  way
     modified and this statement is not removed.  Usage  by  and for commercial
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
     or send an email to license@isb-sib.ch).
     --------------------------------------------------------------------------
     EMBL; AE005577; AAG58658.1; -.
     EMBL; AP002565; BAB37820.1; -.
     PIR; E91178; E91178.
     PIR; F86024; F86024.
     InterPro; IPR002129; Pyridoxal_deC.
     Pfam; PF00282; pyridoxal_deC; 1.
     PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
     Lyase; Decarboxylase; Pyridoxal phosphate; Multigene family;
     Complete proteome.
     BINDING     276    276       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
     SEQUENCE   466 AA;  52699 MW;  C7F9623DBB24E489 CRC64;

 Query Match             38.6%;  Score 1009;  DB 1;  Length 466;
 Best Local Similarity   44.7%;  Pred. No. 3.4e-63;
 Matches  196;  Conservative   84;  Mismatches  152;  Indels    6;  Gaps    3;

y            16 STFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKL 75
                | | :: : |     : | : :  : |:||||||| |||| | |||:|  || :    ||
b            16 SRFGAKAISTIAESKRFPLHEMRDDVAFQIINDELYLDGNARQNLATFCQTWDDENVHKL 75

y            76 IMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGL 135
                :  |||||::| :||| :  :  |||||:| |::||  :   |||  |:||||| || |:
b            76 MDLSINKNWIDKEEYPQSAAIDLRCVNMVADLWHAPAPKNGQAVGTNTIGSSEACMLGGM 135

y           136 AFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
                | | :|: : :| ||| ||||:| |  ||:|| |||||::|||:|: :  |   |||::
b           136 AMKWRWRKRMEAAGKPTDKPNLVCGP-VQICWHKFARYWDVELREIPMRPGQLFMDPKRM 194

y           196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
                ::  ||||| |    | |  | :|  : |:| | :   :|| |  :|:|||||||:|||:
b           195 IEACDENTIGVVPTFGVTYTGNYEFPQPLHDALDKFQADTGIDIDMHIDAASGGFLAPFV 254

y           256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
                 |:: |||||| ||||: ||||:||   | ||||||::| ||:||:|:::|||    ||
b           255 APDIVWDFRLPRVKSISASGHKFGLAPLGCGWVIWRDEEALPQELVFNVDYLGGQIGTFA 314

y           316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIV---SKDEG 372
                :|||: : |||||||: :||| |||  |     :    | : : |   :  :     |||
b           315 INFSRPAGQVIAQYYEFLRLGREGYTKVQNASYQVAAYLADEIAKLGPYEFICTGRPDEG 374

y           373 VPLVAFSLKDSS--CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLA 430
                :| | | ||:     :| :::|: ||  || |||:|:   |  | |:|:: |  |    |
b           375 IPAVCFKLKEGEDPGYTLYDLSERLRLRGWQVPAFTLGGEATDIVVMRIMCRRGFEMDFA 434

y           431 ERLVIDIEKVMRELDELP 448
                | |: | :  :: | : |
```

```
Db        435 ELLLEDYKASLKYLSDHP 452

RESULT 10
DCEC_LISIN
ID   DCEC_LISIN     STANDARD;      PRT;   467 AA.
AC   Q928K4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probabl glutamate decarboxylase gamma (EC 4.1.1.15) (GAD-gamma).
GN   LIN2528.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / Serovar 6a;
RX   MEDLINE=21537279; PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
CC   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
CC   -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
CC       GAD, HDC AND TYRDC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AL596172; CAC97755.1; -.
DR   PIR; AC1748; AC1748.
DR   ListiList; LIN02528; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
DR   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
KW   Lyase; Decarboxylase; Pyridoxal phosphate; Complete proteome.
FT   BINDING     278    278       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   467 AA;  53690 MW;  008AB18F6E274590 CRC64;

  Query Match             36.8%;  Score 963;  DB 1;  Length 467;
  Best Local Similarity   39.9%;  Pred. No. 5.6e-60;
  Matches  186;  Conservative   98;  Mismatches  166;  Indels   16;  Gaps    6;

Qy          8 SESDVSVHST-----FASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLAS 62
              ||:|   | :     | |    ||:|:: : :  :    |||:: |:|| :|| | |||:
Db          4 SENDKRKHESYRIPLFGSEEESTSIPKYVLKKEPMEPRIAYQLVKDQLMDEGNARQNLAT 63

Qy         63 FVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVG 122
              |  |:|| | : |:  :: || :|  ||| | ||:|||||::| |:||| |   : :|
Db         64 FCQTYMEKEAEILMAETLEKNAIDKSEYPQTAELENRCVNILADLWNAPKE--MSYLGTS 121

Qy        123 TVGSSEAIMLAGLAFKRKWQNKRKAEGKPVD--KPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||| || ||| | :|:|  :  |  :   :||::  :  |||||||  |::|:::
Db        122 TVGSSEACMLGGLAMKFRWRNNAEKRGLDIQAKRPNLIISSGYQVCWEKFCVYWDVDMRV 181

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              | : : :  :| ::  ::||| || :  ||| |  |:|:|: ||:: :   |:
```

```
b         182 VPMDKEHLSLDVEKVFELVDEYTIGIVGILGITYTGKFDDIALLDEKVEAYNEANEHQLV 241

y         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||:| |||    ||: ||| |||||  | ||| ||||||||| |:||::|::|| ||:||
b         242 IHIDGASGAMFTPFVNPELPWDFRLKNVVSINTSGHKYGLVYPGVGWILWKDKEYLPKEL 301

y         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              || ::|||   ||  :|||: :||:| |||  :| | |||| : |  ::  : | : :||
b         302 IFEVSYLGGSMPTMAINFSRSASQIIGQYYNFLRYGFEGYREIHEKTKKTALYLSKTVEK 361

y         361 TERFNIVSKDEGVPLVAFSLKDSSC--HTEFEISDMLRRYGWIVPAYTMPPNAQHITVLR 418
              :  | |::    :|:| : |||      | ::::| |   || |||| :| :     : |
b         362 SGYFEIINDGSNLPIVCYKLKDDLDVEWTLYDLADQLLMKGWQVPAYPLPADLSDTIIQR 421

y         419 VVIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSES 464
               | | |    :||    |    :  |:   :||::     :|:::|
b         422 FVCRADLGYNVAEEFAADFADALHNLEH--ARVLYH---DKERNDS 462

ESULT 11
CEA_LISMO
D   DCEA_LISMO     STANDARD;      PRT;   462 AA.
C   Q9F5P3; Q8Y9S6;
T   28-FEB-2003 (Rel. 41, Created)
T   28-FEB-2003 (Rel. 41, Last sequence update)
T   28-FEB-2003 (Rel. 41, Last annotation update)
E   Glutamate decarboxylase alpha (EC 4.1.1.15) (GAD-alpha).
N   GADA OR LMO0447.
S   Listeria monocytogenes.
C   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
X   NCBI_TaxID=1639;
N   [1]
P   SEQUENCE FROM N.A.
C   STRAIN=LO28 / Serovar 1/2c;
X   MEDLINE=21206233; PubMed=11309128;
A   Cotter P.D., Gahan C.G.M., Hill C.;
T   "A glutamate decarboxylase system protects Listeria monocytogenes in
T   gastric fluid.";
L   Mol. Microbiol. 40:465-475(2001).
N   [2]
P   SEQUENCE FROM N.A.
C   STRAIN=EGD-e / Serovar 1/2a;
X   MEDLINE=21537279; PubMed=11679669;
A   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
A   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
A   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
A   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
A   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
A   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
A   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
A   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
A   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
A   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
A   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
T   "Comparative genomics of Listeria species.";
L   Science 294:849-852(2001).
C   -!- FUNCTION: Converts internalized glutamate to GABA and increases
C       the internal pH. Involved in glutamate-dependent acid resistance
C       in gastric fluid.
C   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
C   -!- COFACTOR: Pyridoxal phosphate (By similarity).
C   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
C       GAD, HDC AND TYRDC).
C   --------------------------------------------------------------------------
C   This SWISS-PROT entry is copyright. It is produced through a collaboration
C   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
C   the European Bioinformatics Institute.  There are no  restrictions on  its
C   use  by  non-profit  institutions as long  as its content  is  in  no way
C   modified and this statement is not removed.  Usage  by  and for commercial
C   entities requires a license agreement (See http://www.isb-sib.ch/announce/
C   or send an email to license@isb-sib.ch).
C   --------------------------------------------------------------------------
R   EMBL; AF309076; AAG22560.1; -.
```

```
DR   EMBL; AL591975; CAC98526.1; -.
DR   PIR; AH1130; AH1130.
DR   ListiList; LMO00447; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
DR   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
KW   Lyase; Decarboxylase; Pyridoxal phosphate; Complete proteome.
FT   BINDING     273    273       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT   CONFLICT    152    152       S -> N (IN REF. 1).
SQ   SEQUENCE   462 AA;  52498 MW;  EA1A442E3E1CE2FA CRC64;

  Query Match             36.5%;  Score 955.5;  DB 1;  Length 462;
  Best Local Similarity   41.3%;  Pred. No. 1.8e-59;
  Matches  183;  Conservative   94;  Mismatches  157;  Indels    9;  Gaps    4;

Qy          7 VSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTT 66
              | :::| |  :| |      ||  :| : |:    |||:: |:|: :|: | |||:|  |
Db          6 VEQNNVPVFGSFES---GQDLPEKRMNKESVDPRIAYQLVKDQLIDEGSARQNLATFCQT 62

Qy         67 WMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGS 126
              :|||| ::::  :: || :|  ||| | :|:: ||||:| |:|  ::|:|  :|  ||||
Db         63 YMEPEAEQIMAETMEKNAIDKSEYPQTAKLESSCVNMLADLWN--VDESEHYMGTSTVGS 120

Qy        127 SEAIMLAGLAFKRKWQNKRKAEGKPV--DKPNIVTGANVQVCWEKFARYFEVELKEVKLS 184
              ||| || |:| | :|::     |  :   ||::|  :  |||||||  |:::||:|| :|
Db        121 SEACMLGGMAMKFRWRSAALKNGLDIHAKKPSLVISSGYQVCWEKFCVYWDIELREVPMS 180

Qy        185 EGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVD 244
              | :  ::    :| ||| || :  ||| |  |:|:|:  ||||: : |     :  ||||
Db        181 EEHLSINTDIIMDYVDEYTIGIVGILGITYTGKFDDIMTLNDLVEDYNNTHDNEVVIHVD 240

Qy        245 AASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHI 304
               |||    ||: | |||||||| | ||| ||||||||| |:||::||:|| |||||:| :
Db        241 GASGAMFTPFVEPGLEWDFRLPNVVSINTSGHKYGLVYPGVGWILWRDKEYLPEELVFDV 300

Qy        305 NYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERF 364
              :|||   ||  :|||: :||:| |||  :| |:|||| :    |:  : | : : :|  |
Db        301 SYLGGHMPTMAINFSRSASQIIGQYYNFLRFGYEGYRQIHMRTRDGALQLSQAVAETGLF 360

Qy        365 NIVSKDEGVPLVAFSLKDSS--CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIR 422
               | :    :|:| : ||| :    | ::::| |:  || |||| :|    :  : | | |
Db        361 EIYNDGANLPIVCYKLKDDANVAWTLYDLADRLQMKGWQVPAYPLPKEMGNTIIQRYVCR 420

Qy        423 EDFSRTLAERLVIDIEKVMRELD 445
               |  : :      |: : : ||:
Db        421 GDLGQNMVTAFKNDLSESIEELN 443


RESULT 12
DCEC_LISMO
ID   DCEC_LISMO     STANDARD;      PRT;   467 AA.
AC   Q8Y4K4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probabl glutamate decarboxylase gamma (EC 4.1.1.15) (GAD-gamma).
GN   LMO2434.
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGD-e / Serovar 1/2a;
RX   MEDLINE=21537279; PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
```

```
A   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
A   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
T   "Comparative genomics of Listeria species.";
L   Science 294:849-852(2001).
C   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
C   -!- COFACTOR: Pyridoxal phosphate (By similarity).
C   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
C       GAD, HDC AND TYRDC).
C   --------------------------------------------------------------------------
C   This SWISS-PROT entry is copyright. It is produced through a collaboration
C   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
C   the European Bioinformatics Institute.  There are no  restrictions on  its
C   use  by  non-profit  institutions as long  as its content  is  in  no  way
C   modified and this statement is not removed.  Usage  by  and for commercial
C   entities requires a license agreement (See http://www.isb-sib.ch/announce/
C   or send an email to license@isb-sib.ch).
C   --------------------------------------------------------------------------
R   EMBL; AL591983; CAD00512.1; -.
R   PIR; AB1379; AB1379.
R   ListiList; LMO02434; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
R   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
W   Lyase; Decarboxylase; Pyridoxal phosphate; Complete proteome.
T   BINDING     278    278       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
Q   SEQUENCE   467 AA;  53640 MW;  64A8CB6FDC82B8CF CRC64;

 Query Match             36.5%;  Score 953.5;  DB 1;  Length 467;
 Best Local Similarity   40.4%;  Pred. No. 2.6e-59;
 Matches  182;  Conservative   96;  Mismatches  162;  Indels   11;  Gaps    5;

y         18 FASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIM 77
             | |    ||:|:: : :  :    |||:: |:|| :|| | |||:|  |:|| | : |:
b         19 FGSEEESTSIPKYVLKKEPMEPRIAYQLVKDQLMDEGNARQNLATFCQTYMEKEAEILMA 78

y         78 SSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
              :: || :|  ||| | ||:|||||::| |:||| :   : :|  ||||||| || |||
b         79 ETLEKNAIDKSEYPQTAELENRCVNILADLWNAPKD--MSYLGTSTVGSSEACMLGGLAM 136

y        138 KRKWQNKRKAEGKPVD--KPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
             | :|:|  :  |  :   :||::  :  |||||  |::|::: | : : :  :|  :
b        137 KFRWRNNAEKRGLDIQAKRPNLIISSGYQVCWEKFCVYWDVDMRVVPMDKNHLSLDVDKV 196

y        196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
              |:||| || |  ||| |  |:|:|::||:: :   |:       ||:| |||    ||:
b        197 FDLVDEYTIGVVGILGITYTGKFDDIQLLDEKVEAYNETNEHQLVIHIDGASGAMFTPFV 256

y        256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
              ||| |||||  |  ||| ||||||||| |:||::|::||  ||:|||| ::|||   ||
b        257 NPELPWDFRLKNVVSINTSGHKYGLVYPGVGWILWKDKEYLPKELIFEVSYLGGSMPTMA 316

y        316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPL 375
             :|||: :||:|| |||  :|| | |||| : |  ::  | | : :||:  | |::    :|:
b        317 INFSRSASQIIGQYYNFLRYGFEGYREIHEKTKKTAIYLAKTVEKSGYFEIINDGANLPI 376

y        376 VAFSLKDSSC--HTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERL 433
             | : :|:      | ::::| |   || |||| :| :     : | | | |    :||
b        377 VCYKMKEGLDVEWTLYDLADQLLMKGWQVPAYPLPADLSDTIIQRFVCRADLGYNVAEEF 436

y        434 VIDIEKVMRELDELPSRVIHKISLGQEKSES 464
               |    :  |:   :||::     :|:::|
b        437 AADFADAIHNLEH--ARVLYH---DKERNDS 462


ESULT 13
CEB_LISIN
D   DCEB_LISIN     STANDARD;      PRT;   464 AA.
C   Q928R9;
T   28-FEB-2003 (Rel. 41, Created)
T   28-FEB-2003 (Rel. 41, Last sequence update)
T   28-FEB-2003 (Rel. 41, Last annotation update)
E   Glutamate decarboxylase beta (EC 4.1.1.15) (GAD-beta).
N   GADB OR LIN2463.
```

```
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / Serovar 6a;
RX   MEDLINE=21537279; PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
CC   -!- FUNCTION: Converts internalized glutamate to GABA and increases
CC       the internal pH. Involved in glutamate-dependent acid resistance
CC       in gastric fluid (By similarity).
CC   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
CC   -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
CC       GAD, HDC AND TYRDC).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AL596172; CAC97690.1; -.
DR   PIR; AB1740; AB1740.
DR   ListiList; LIN02463; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
DR   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
KW   Lyase; Decarboxylase; Pyridoxal phosphate; Complete proteome.
FT   BINDING     275    275       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   464 AA;  53599 MW;  4C35CD1395ADF481 CRC64;

  Query Match             36.3%;  Score 948;  DB 1;  Length 464;
  Best Local Similarity   41.7%;  Pred. No. 6.2e-59;
  Matches  180;  Conservative   92;  Mismatches  154;  Indels    6;  Gaps    3;

Qy         18 FASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIM 77
              | |      :|:: : :  :    ||::: |||: :|: | |||:|  |:|| | ||:
Db         16 FGSSAEDRDIPKYTLAKEPLEPRIAYRLVKDELLDEGSARQNLATFCQTYMEDEATKLMS 75

Qy         78 SSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
               :: || :|  ||| | ||:|||||:|| |::|| :  :  :|  |:||||| || |:|
Db         76 ETLEKNAIDKSEYPRTAELENRCVNIIADLWHAPKD--QKFMGTSTIGSSEACMLGGMAM 133

Qy        138 KRKWQNKRKAEGKPV--DKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
              |  |: : :  |  :   |||:|  :  |||||||  |:::::: | : : :  ::  |
Db        134 KFAWRKRAEKLGLDIYAQKPNLVISSGYQVCWEKFCVYWDIDMRVVPMDKDHMQLNTDQV 193

Qy        196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
              :| ||| || |  ||| |  | ::|:  ||: | | | :| :   ||||||||||  ||:
Db        194 LDYVDEYTIGVVGILGITYTGRYDDIYALNEKLEEYNSKTDYKVYIHVDAASGGFFTPFV 253

Qy        256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
               |:: |||||  |  ||| ||||||||| |||||:|:::  ||||||| ::||| : ||
Db        254 EPDIIWDFRLKNVISINTSGHKYGLVYPGIGWVLWKDESYLPEELIFKVSYLGGEMPTMQ 313

Qy        316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPL 375
              :|||: :| :| |||  :| | |||| : :   :    |   :|:|  |:| :    :|:
Db        314 INFSRSASHIIGQYYNFLRYGFEGYRTIHQKTSDVAQYLAHAVEQTGYFDIYNDGSHLPI 373
```

```
/        376 VAFSLKDSS--CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERL 433
             | : ||| :    | ::::| |:  || |||| :| | ::| : | | | |    :||
ɔ        374 VCYKLKDDANVKWTLYDLADRLQMRGWQVPAYPLPKNLENIIIQRYVCRADLGFNMAEEF 433

/        434 VIDIEKVMRELD 445
             : | :  ::||:
ɔ        434 IQDFQASIQELN 445

ESULT 14
CEB_LISMO
D   DCEB_LISMO     STANDARD;      PRT;   464 AA.
C   Q9EYW9; Q8Y4S0; Q9AGQ0;
T   28-FEB-2003 (Rel. 41, Created)
T   28-FEB-2003 (Rel. 41, Last sequence update)
T   28-FEB-2003 (Rel. 41, Last annotation update)
E   Glutamate decarboxylase beta (EC 4.1.1.15) (GAD-beta).
N   GADB OR LMO2363.
S   Listeria monocytogenes.
C   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
X   NCBI_TaxID=1639;
N   [1]
P   SEQUENCE FROM N.A.
C   STRAIN=LO28 / Serovar 1/2c, and EGD5;
X   MEDLINE=21206233; PubMed=11309128;
A   Cotter P.D., Gahan C.G.M., Hill C.;
T   "A glutamate decarboxylase system protects Listeria monocytogenes in
T   gastric fluid.";
L   Mol. Microbiol. 40:465-475(2001).
N   [2]
P   SEQUENCE FROM N.A.
C   STRAIN=EGD-e / Serovar 1/2a;
X   MEDLINE=21537279; PubMed=11679669;
A   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
A   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
A   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
A   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
A   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
A   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
A   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
A   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
A   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
A   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
A   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
T   "Comparative genomics of Listeria species.";
L   Science 294:849-852(2001).
C   -!- FUNCTION: Converts internalized glutamate to GABA and increases
C       the internal pH. Involved in glutamate-dependent acid resistance
C       in gastric fluid.
C   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
C   -!- COFACTOR: Pyridoxal phosphate (By similarity).
C   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
C       GAD, HDC AND TYRDC).
C   --------------------------------------------------------------------------
C   This SWISS-PROT entry is copyright. It is produced through a collaboration
C   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
C   the European Bioinformatics Institute.  There are no  restrictions on  its
C   use  by  non-profit  institutions as long  as its content  is  in  no  way
C   modified and this statement is not removed.  Usage  by  and for commercial
C   entities requires a license agreement (See http://www.isb-sib.ch/announce/
C   or send an email to license@isb-sib.ch).
C   --------------------------------------------------------------------------
R   EMBL; AF309077; AAG22562.1; -.
R   EMBL; AF329447; AAK17187.1; -.
R   EMBL; AL591983; CAD00441.1; -.
R   PIR; AC1370; AC1370.
R   ListiList; LMO02363; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
R   PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
W   Lyase; Decarboxylase; Pyridoxal phosphate; Complete proteome.
T   BINDING     275    275       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
```

```
FT   VARIANT      92     92       A -> P (IN STRAIN LO28).
FT   VARIANT     124    124       E -> D (IN STRAIN LO28).
FT   VARIANT     261    261       F -> L (IN STRAIN EGD5).
FT   VARIANT     375    375       C -> R (IN STRAIN EGD5).
FT   VARIANT     380    381       DD -> TT (IN STRAIN LO28).
SQ   SEQUENCE   464 AA;  53542 MW;  F2E2778CFD1E2C36 CRC64;

  Query Match             35.9%;  Score 940;  DB 1;  Length 464;
  Best Local Similarity   41.4%;  Pred. No. 2.2e-58;
  Matches  179;  Conservative   93;  Mismatches  154;  Indels    6;  Gaps    3;

Qy         18 FASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIM 77
              | |      :|:: : :  :    ||::: |||: :|: | |||:|  |:|| |  ||:
Db         16 FGSSAEDRDIPKYTLGKEPLEPRIAYRLVKDELLDEGSARQNLATFCQTYMEDEATKLMS 75

Qy         78 SSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
               :: || :|  ||| | ||:|||||:|| |::|| :  :  :|  |:||||| || |:|
Db         76 ETLEKNAIDKSEYPRTAELENRCVNIIADLWHAPKD--QKFMGTSTIGSSEACMLGGMAM 133

Qy        138 KRKWQNKRKAEGKPV--DKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
              |  |: : :  |  :   |||:|  :  |||||||  |:::::: | : : :  ::  |
Db        134 KFAWRKRAEKLGLDIYAKKPNLVISSGYQVCWEKFCVYWDIDMRVVPMDKEHMQLNTDQV 193

Qy        196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
              :| ||| || |  ||| |  | ::|:  ||: | | | :| :   ||||||||||  ||:
Db        194 LDYVDEYTIGVVGILGITYTGRYDDIYALNEKLEEYNSKTDYKVYIHVDAASGGFFTPFV 253

Qy        256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
               |:: |||||  | ||| ||||||||| |||||:|:::  ||||||| ::||| : ||
Db        254 EPDIIWDFRLKNVISINTSGHKYGLVYPGIGWVLWKDESYLPEELIFKVSYLGGEMPTMQ 313

Qy        316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPL 375
              :|||: :| :| |||  :| | |||| : :   :    |   :|:|  |:| :    :|:
Db        314 INFSRSASHIIGQYYNFLRYGFEGYRTIHQKTSDVAQYLAHAVEQTGYFDIFNDGSHLPI 373

Qy        376 VAFSLKDSS--CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERL 433
              | : ||| :    | ::::| |:  || |||| :| : ::| : | | | |    :||
Db        374 VCYKLKDDANVNWTLYDLADRLQMRGWQVPAYPLPKSLENIIIQRYVCRADLGFNMAEEF 433

Qy        434 VIDIEKVMRELD 445
              : | :  ::||:
Db        434 IQDFQASIQELN 445

RESULT 15
DCE_YEAST
ID   DCE_YEAST      STANDARD;      PRT;   585 AA.
AC   Q04792;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Glutamate decarboxylase (EC 4.1.1.15) (GAD).
GN   GAD1 OR YMR250W OR YM9920.04.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169872;
RA   Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA   Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA   Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA   Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT   XIII.";
RL   Nature 387:90-93(1997).
CC   -!- CATALYTIC ACTIVITY: L-glutamate = 4-aminobutanoate + CO(2).
CC   -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE GROUP II DECARBOXYLASE FAMILY (DDC,
CC       GAD, HDC AND TYRDC).
CC   --------------------------------------------------------------------------
```

```
   This SWISS-PROT entry is copyright. It is produced through a collaboration
   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
   the European Bioinformatics Institute.  There are no  restrictions on  its
   use  by  non-profit  institutions as long  as its content  is  in  no  way
C  modified and this statement is not removed.  Usage  by  and for commercial
C  entities requires a license agreement (See http://www.isb-sib.ch/announce/
C  or send an email to license@isb-sib.ch).
C  -----------------------------------------------------------------------
R  EMBL; Z48639; CAA88577.1; -.
R  PIR; S53072; S53072.
R  SGD; S0004862; GAD1.
R  GO; GO:0005737; C:cytoplasm; IDA.
R  GO; GO:0004351; F:glutamate decarboxylase activity; IPI.
R  GO; GO:0006538; P:glutamate catabolism; IPI.
R  GO; GO:0006979; P:response to oxidative stress; IMP.
R  InterPro; IPR002129; Pyridoxal_deC.
R  Pfam; PF00282; pyridoxal_deC; 1.
R  PROSITE; PS00392; DDC_GAD_HDC_YDC; FALSE_NEG.
W  Lyase; Decarboxylase; Pyridoxal phosphate.
T  BINDING     318    318       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
Q  SEQUENCE   585 AA;  65990 MW;  DDD22868CEB7B955 CRC64;

Query Match             33.9%;  Score 887.5;  DB 1;  Length 585;
Best Local Similarity   39.2%;  Pred. No. 1.4e-54;
Matches  200;  Conservative   78;  Mismatches  175;  Indels   57;  Gaps   12;

y         29 RFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDMD 88
             :: :|:  :| : :||:|::|| ||||| ||||||| |:   :  |||  :: ||  | |
o         59 KYSVPKKGLPADLSYQLIHNELTLDGNPHLNLASFVNTFTTDQARKLIDENLTKNLADND 118

y         89 EYPVTTELQNRCVNMIAHLFNA-PLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKA 147
             |||   ||  ||::|:| |::| | ||    :|  | |||||||| ||| |::|::: |
o        119 EYPQLIELTQRCISMLAQLWHANPDEE---PIGCATTGSSEAIMLGGLAMKKRWEHRMKN 175

y        148 EGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLS-EGYYVMDPQQAVDMVDENTICV 206
              ||   ||||:  :  ||  ||| |||||| : | :|   ::::||:   | ||||||
o        176 AGKDASKPNIIMSSACQVALEKFTRYFEVECRLVPVSHRSHHMLDPESLWDYVDENTIGC 235

y        207 ADILGSTLNGEFEDVKLLNDLLVE-KNKETGW---DTPIHVDAASGGFIAPF-------- 254
               |||:|  |  |:|: : |:| : : |   |   | ||| | |||||| ||
o        236 FVILGTTYTGHLENVEKVADVLSQIEAKHPDWSNTDIPIHADGASGGFIIPFGFEKEHMK 295

y        255 LYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTF 314
              |    | |  | | |:| ||||:||   |:|||:||::  | :|| | : |||  : ||
b        296 AYGMERWGFNHPRVVSMNTSGHKFGLTTPGLGWVLWRDESLLADELRFKLKYLGGVEETF 355

y        315 TLNFSKGSSQVIAQYYQLIRLGHEGYRNVMEN--------------------CRENMIVL 354
              ||||:   ||: ||:  : ||| |||   :|                    | | :  :
b        356 GLNFSRPGFQVVHQYFNFVSLGHSGYRTQFQNSLFVARAFSFELLNSSKLPGCFEIVSSI 415

y        355 REGLEKTERFNIVSKD---------EGVPLVAFSLKDSSCHTEFE------ISDMLRRYG 399
              | :|       | ||          ||||||| |     | |:       :| :||  |
b        416 HESIENDSAPKSV-KDYWEHPQAYKPGVPLVAFKL-SKKFHEEYPEVPQAILSSLLRGRG 473

y        400 WIVPAYTMP---PNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKIS 456
             ||:| | :|     :    ||||| | :    ||: |::||| :: :|     :| | |
b        474 WIIPNYPLPKATDGSDEKEVLRVVFRSEMKLDLAQLLIVDIESILTKLIHSYEKVCHHIE 533

y        457 LGQEKSESNSDNLMVTVKKSDIDKQRDIIT 486
             |  |::     : :  :  :    | || |
b        534 LASEQTPERKSSFIYEMLLALASPQDDIPT 563

earch completed: October 22, 2003, 15:06:17
ob time : 40 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

M protein - nucleic search, using frame_plus_p2n model

un on:          October 22, 2003, 15:14:55 ; Search time 2444 Seconds
                                           (without alignments)
                                           4992.163 Million cell updates/sec

itle:           US-10-006-852-2
erfect score:   2615
equence:        1 MVLSHAVSESDVSVHSTFAS..........DIITGWKKFVADRKKTSGIC 502

coring table:   BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

earched:        22781392 seqs, 12152238056 residues

otal number of hits satisfying chosen parameters:      45562784

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

ommand line parameters:
MODEL=frame+_p2n.model -DEV=xlp
Q=/cgn2_1/USPTO_spool_p/US10006852/runat_21102003_164104_2461/app_query.fasta_1.647
DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US10006852_@CGN_1_1_3549_@runat_21102003_164104_2461 -NCPU=6 -ICPU=3
NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :       EST:*
                1:  em_estba:*
                2:  em_esthum:*
                3:  em_estin:*
                4:  em_estmu:*
                5:  em_estov:*
                6:  em_estpl:*
                7:  em_estro:*
                8:  em_htc:*
                9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  em_gss_hum:*
                18:  em_gss_inv:*
                19:  em_gss_pln:*
                20:  em_gss_vrt:*
                21:  em_gss_fun:*
                22:  em_gss_mam:*
                23:  em_gss_mus:*
                24:  em_gss_pro:*
                25:  em_gss_rod:*
                26:  em_gss_phg:*
                27:  em_gss_vrl:*
                28:  gb_gss1:*
```

```
                29:  gb_gss2:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                      %
Result               Query
  No.     Score      Match Length DB  ID                     Description
--------------------------------------------------------------------------------
    1      1622      62.0   1168  11  AY103733               AY103733 Zea mays
    2      1332      50.9    811  14  CB655721               CB655721 OSJNEc09G
    3      1281      49.0    821  10  BE040860               BE040860 OF12G05 O
    4      1279      48.9    837  10  BF275438               BF275438 GA__Eb002
    5      1275      48.8    819  13  BQ165651               BQ165651 EST611520
    6      1225      46.8    775  13  BU026594               BU026594 QHG17F24.
    7      1220      46.7    786  14  CA813481               CA813481 CA48LN10I
    8    1217.5      46.6    823  10  BF263784               BF263784 HV_CEa000
    9      1203      46.0    762  13  BQ971999               BQ971999 QHB9G09.y
   10      1196      45.7    803  10  BE054156               BE054156 GA__Ea003
   11      1189      45.5    750  13  BU023643               BU023643 QHF12C03.
   12    1187.5      45.4    786  13  BU027739               BU027739 QHG7H16.y
   13      1185      45.3    746  13  BU025264               BU025264 QHF8J20.y
   14      1185      45.3    746  13  BU027680               BU027680 QHG7B13.y
   15      1184      45.3    746  13  BU025649               BU025649 QHG10J06.
   16      1182      45.2    752  13  BU026039               BU026039 QHG12O09.
   17      1182      45.2    769  13  BU025918               BU025918 QHG12C07.
   18      1180      45.1    744  13  BU025839               BU025839 QHG11L03.
   19      1179      45.1    744  13  BU027771               BU027771 QHG7K14.y
   20      1178      45.0    744  13  BU025655               BU025655 QHG10K09.
   21      1178      45.0    744  13  BU025893               BU025893 QHG11P15.
   22    1177.5      45.0    808  14  CB893221               CB893221 EST646013
   23      1175      44.9    744  13  BU026030               BU026030 QHG12N09.
   24      1175      44.9    745  13  BU025328               BU025328 QHF8O15.y
   25      1171      44.8    752  13  BQ841623               BQ841623 WHE2983_B
   26      1168      44.7    698  14  CD484071               CD484071 atr01-6ms
   27      1168      44.7    746  13  BU027486               BU027486 QHG5P23.y
   28      1167      44.6    790  10  BG647367               BG647367 EST508986
   29      1164      44.5    744  13  BU025929               BU025929 QHG12D11.
   30      1162      44.4    732  13  BU026249               BU026249 QHG14D05.
   31    1161.5      44.4    781  12  BI308490               BI308490 EST529900
   32      1160      44.4    781  13  BU025900               BU025900 QHG12A05.
   33      1159      44.3    733  10  BG594946               BG594946 EST493624
   34      1154      44.1    721  13  BU027749               BU027749 QHG7I10.y
   35      1154      44.1    746  13  BU026554               BU026554 QHG17D12.
   36      1152      44.1    726  13  BU024451               BU024451 QHF3g03.y
   37      1152      44.1    726  13  BU027466               BU027466 QHG5N09.y
   38      1152      44.1    742  13  BQ916451               BQ916451 QHB17O10.
   39      1151      44.0    793  14  CB341828               CB341828 CA32EN000
   40      1150      44.0    742  13  BU025333               BU025333 QHF8P03.y
   41      1146      43.8    727  13  BU023920               BU023920 QHF14N08.
   42      1144      43.7    716  13  BU027449               BU027449 QHG5L24.y
   43      1143      43.7    715  13  BU025894               BU025894 QHG11P17.
   44      1142      43.7    721  10  BG589644               BG589644 EST497486
   45      1142      43.7    743  13  BU025776               BU025776 QHG11F24.


                                  ALIGNMENTS

RESULT 1
AY103733
LOCUS       AY103733                1168 bp    mRNA    linear   HTC 16-OCT-2002
DEFINITION  Zea mays PCO153431 mRNA sequence.
ACCESSION   AY103733
VERSION     AY103733.1  GI:21206811
KEYWORDS    HTC.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1168)
```

```
 AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
           Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
           Overgo Probes
 JOURNAL   Unpublished (2002)
EFERENCE   2  (bases 1 to 1168)
 AUTHORS   Coe,E.H.
 TITLE     Direct Submission
 JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
           Missouri, Columbia, MO 65211, USA
OMMENT     If you are interested in getting corresponding physical clones,
           these are publicly available from ZmDB and may be found by BLAST
           searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
           www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
           maize cDNA sequences is either Virginia Walbot, Stanford or Pat
           Schnable, Iowa State, then clones may be requested from ZmDB:
           www.zmdb.iastate.edu.
EATURES             Location/Qualifiers
    source          1..1168
                    /organism="Zea mays"
                    /mol_type="mRNA"
                    /db_xref="MaizeDB:638900"
                    /db_xref="taxon:4577"
                    /clone_lib="Maize Mapping Project/DuPont Cornsensus
                    Library"
                    /note="this sequence is part of a project of EST
                    assemblies resulting from the application of public
                    contigs to seed DuPont contigs; this resource was
                    assembled by DuPont as part of a collaboration for the
                    overgo addressing of BACs in conjunction with the Maize
                    Mapping Project"
ASE COUNT      275 a    332 c    345 g    216 t
RIGIN

lignment Scores:
red. No.:              6.96e-178      Length:         1168
core:                  1622.00        Matches:        309
ercent Similarity:     92.90%         Conservative:   18
est Local Similarity:  87.78%         Mismatches:     23
uery Match:            62.03%         Indels:         3
B:                     11             Gaps:           1

S-10-006-852-2 (1-502) x AY103733 (1-1168)

y           1 MetValLeuSerHisAlaValSer------GluSerAspValSerValHisSerThrPhe 18
              ||||||||||||||||||   |||      :::   |||   :::|||   |||||||||
b         107 ATGGTGCTCTCGCACGCGAGCTCCGGCCGGGACGACGACGACGCGGTGGCCTCCACCTTC 166

y          19 AlaSerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIlePro 38
              ||||||||||||||||||      |||||||||::::::|||||||||   |||||||||
b         167 GCGTCGCGCTACGTGCGCGAGCGCCTGCCGCGGTACCGGATGCCGGAGCGGTCGATCCCG 226

y          39 LysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeu 58
              :::|||||||||   |||||||||:::|||||||||||||||||||||||||||||||||
b         227 CGGGAGGCGGCGCAGCAGATCATCAGCGACGAGCTCATGCTGGACGGCAACCCGCGCCTG 286

y          59 AsnLeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSer 78
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         287 AACCTGGCGTCCTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATCATGGAC 346

Qy         79 SerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsn 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 TCCATCAACAAGAACTACGTCGACATGGACGAGTACCCCGTCACCACTGAGCTCCAGAAC 406

Qy         99 ArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAla 118
              |||||||||||||||||||||||||||||||||||||||::::::|||   |||||||||
Db        407 CGCTGCGTAAACATGATAGCTCACCTGTTCAACGCGCCGATCAAGGAGGACGAGACAGCT 466

Qy        119 ValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLys 138
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 ATTGGGGTCGGAACAGTGGGATCCTCGGAAGCGATCATGCTTGCAGGTCTGGCGTTCAAG 526

Qy        139 ArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleVal 158
```

```
              ||||||||||||||||||||||||   :::|||||||||   |||:::||||||||||||
Db        527 AGGAAATGGCAGAACAAGAGGAAGGAGCAGGGGAAACCGTGTGACAGACCCAACATCGTC 586

Qy        159 ThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeu 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 ACCGGTGCCAACGTTCAGGTCTGCTGGGAGAAGTTTGCAAGGTATTTTGAGGTGGAGCTG 646

Qy        179 LysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMet 198
              ||||||||||||||||||||||||||||||||||||||||||   :::||||||:::|||
Db        647 AAGGAGGTTAAGCTATCAGAAGGGTACTACGTCATGGACCCCGTCAAGGCTGTTGAGATG 706

Qy        199 ValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPhe 218
              |||||||||||||||||||||||||||   ||||||||||||||||||   |||||||||
Db        707 GTGGACGAGAACACTATCTGTGTCGCGGCGATCTTGGGGTCGACTCTCACCGGAGAGTTC 766

Qy        219 GluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAsp 238
              ||||||||||||   |||||||||||||||   |||||||||||||||||||||||||||
Db        767 GAGGATGTCAAGCAGTTGAACGACCTTCTTACAGAGAAGAACAAGGAAACTGGGTGGGAC 826

Qy        239 ThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGlu 258
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GTGCCTATCCACGTCGACGCCGCGAGCGGAGGGTTCATAGCGCCTTTCCTCTACCCTGAG 886

Qy        259 LeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyr 278
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 CTCGAGTGGGACTTCAGGCTGCCTCTGGTGAAGAGCATCAATGTCAGCGGGCACAAGTAC 946

Qy        279 GlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGlu 298
              ||||||||||||   |||:::||||||||||||||||||   |||:::||||||||||||
Db        947 GGCCTCGTCTACCCTGGTGTTGGCTGGGTCATCTGGCGGAC-AAGGATGATCTGCCTGAG 1005

Qy        299 GluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPhe 318
              ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db       1006 GAGCTCATCTTCCATATCAACTACCTGGGGACGGACCAGCCTACGTTCACGCTCAACTTC 1065

Qy        319 SerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGlu 338
              ||||||||||||   |||:::|||||||||||||||||||||||||||||||||   |||
Db       1066 TCCAAAGGTTCCTGTCAGATCATTGCGCAGTACTATCAACTCATTCGCCTAGGCTTCGAG 1125

Qy        339 GlyTyrArgAsnValMetGluAsnCysArgGluAsn 350
              ||||||||||||:::||||||||||||:::||||||
Db       1126 GGTACAGAAACATCATGGAGAACTGCCAGGAGAAC 1161

RESULT 2
CB655721
LOCUS       CB655721                 811 bp    mRNA    linear   EST 09-APR-2003
DEFINITION  OSJNEc09G09.f OSJNEc Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEc09G09 5', mRNA sequence.
ACCESSION   CB655721
VERSION     CB655721.1  GI:29659446
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 811)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR PRimers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
```

```
           Plate: 09  row: G  column: 09
           Seq primer: gta aaa cga cgg cca gtg.
EATURES              Location/Qualifiers
    source           1. .811
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEc09G09"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEc"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 6 hrs after innoculation with Rice Blast (C9240-1)"
ASE COUNT      182 a    236 c    244 g    149 t
RIGIN

lignment Scores:
red. No.:             2.52e-144      Length:        811
core:                 1332.00        Matches:       249
ercent Similarity:    95.91%         Conservative:  9
est Local Similarity: 92.57%         Mismatches:    11
uery Match:           50.94%         Indels:        0
B:                    14             Gaps:          0

S-10-006-852-2 (1-502) x CB655721 (1-811)

y          61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||   :::|||:::
b           3 GCGTCGTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATCCAGGCCTCCGTC 62

y          81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          63 AACAAGAACTACGTCGACATGGACGAGTACCCCGTCACCACCGAACTCCAGAACCGATGT 122

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||||||||||||||||||||||||||||||||   ::::::|||||||||||||||
Db        123 GTGAACATGATTGCACACCTCTTCAATGCTCCTCTAGGGGACTCTGAAACGGCCGTCGGA 182

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        183 GTCGGCACTGTCGGCTCGTCTGAGGCCATCATGCTCGCCGGTTTGGCCTTCAAGAGGAGG 242

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||   |||||||||   ||||||||||||||||||||||||
Db        243 TGGCAGAACAAGATGAAGGCAGCCGGCAAGCCATGCGACAAGCCTAACATTGTCACCGGC 302

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 GCCAATGTCCAAGTTTGCTGGGAGAAGTTCGCGCGATACTTCGAGGTTGAGCTCAAGGAA 362

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||:::|||||||||||||||||||||   :::||||||||||||||||||
Db        363 GTGAAGCTGAGTGACGGCTACTACGTCATGGACCCAGCTAAGGCCGTGGATATGGTCGAC 422

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        423 GAGAACACCATCTGCGTCGCGGCGATCCTCGGGTCGACGCTGAACGGGGAGTTCGAGGAC 482

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||||||||||||||||||||   :::||||||   |||||||||||||||||||||
Db        483 GTGAAGCTGCTCAACGATCTGCTCACCAAGAAGAACGCTGAAACAGGCTGGGACACGCCG 542

Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 ATCCACGTGGACGCGGCGAGCGGCGGGTTCATCGCGCCGTTCCTGTACCCGGAGCTGGAG 602

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 TGGGACTTCCGGCTGCCGCTGGTGAAGAGCATCAACGTGAGCGGGCACAAGTACGGCCTC 662

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
```

```
              |||||||||||||||||||||   |||||||||:::||||||||||||||||||||||||
Db        663 GTCTACGCCGGGATCGGGTGGTGCATCTGGAGGAGCAAGGAGGATCTGCCTGAGGAGCTC 722

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 ATCTTCCACATCAACTACCTCGGCGCCGACCAGCCCACCTTCACCCTCAACTTCTCCAAG 782

Qy        321 GlySerSerGlnValIleAlaGlnTyr 329
              |||   |||||||||||||||||||||
Db        783 GGGTTCAGCCAGGTCATTGCACAGTAT 809

RESULT 3
BE040860
LOCUS       BE040860                 821 bp    mRNA    linear   EST 07-JUN-2000
DEFINITION  OF12G05 OF Oryza sativa cDNA 5' similar to glutamate decarboxylase
            1, mRNA sequence.
ACCESSION   BE040860
VERSION     BE040860.1  GI:8336367
KEYWORDS    EST.
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 821)
  AUTHORS   Bohnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferrea
            ,H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C.,
            Scara,G., Wheeler,M. and Zepeda,G.R.
  TITLE     Functional Genomics of Plant Stress Tolerance
  JOURNAL   Unpublished
COMMENT     Contact: Michalowski,C.B.
            University of Arizona
            Bio Sciences West room 513, Tucson, AZ 85721, USA
            Tel: 520-621-7982
            Fax: 520-621-1697
            Email: cbm@u.arizona.edu
            An open reading frame exists.
            Insert Length: 1   Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1. .821
                     /organism="Oryza sativa"
                     /mol_type="mRNA"
                     /strain="Pokkali"
                     /db_xref="taxon:4530"
                     /tissue_type="entire plant"
                     /dev_stage="2 weeks"
                     /clone_lib="OF"
                     /note="1 week 150mM NaCl"
BASE COUNT      185 a    231 c    254 g    149 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              2.26e-138      Length:        821
Score:                  1281.00        Matches:       243
Percent Similarity:     94.42%         Conservative:  11
Best Local Similarity:  90.33%         Mismatches:    15
Query Match:            48.99%         Indels:        1
DB:                     10             Gaps:          0

US-10-006-852-2 (1-502) x BE040860 (1-821)

Qy         46 IleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAlaSerPheValThr 65
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 ATCAACGACGAGCTGATGCTGGACGGCAACCCGCGGCTGAACCTCGCGTCGTTCGTCACC 68

Qy         66 ThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIleAsnLysAsnTyrVal 85
              |||||||||||||||||||||||||||||||||   :::|||:::|||||||||||||||
Db         69 ACGTGGATGGAGCCCGAGTGCGACAAGCTCATCCAGGCCTCCGTCAACAAGAACTACGTC 128

Qy         86 AspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIleAla 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 GACATGGACGAGTACCCCGTCACCACCGAGCTCCAGAACCGATGTGTGAACATGATTGCA 188
```

```
y          106 HisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyValGlyThrValGly 125
               |||||||||||||||||||||   ::::::||||||||||||||||||||||||||||||
b          189 CACCTCTTCAATGCTCCTCTAGGGGACTCTGAAACGGCCGTCGGAGTCGGCACTGTCGGC 248

y          126 SerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLysArg 145
               ||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
b          249 TCGTCTGAGGCCATCATGCTCGCCGGTTTGGCCTTCAAGAGGAGGTGGCAGAACAAGATG 308

y          146 LysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGlnVal 165
               ||||||   |||||||||   |||||||||||||||||||||||||||||||||||||||
b          309 AAGGCAGCCGGCAAGCCATGCGACAAGCCTAACATTGTCACCGGCGCCAATGTCCAAGTT 368

y          166 CysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSerGlu 185
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
b          369 TGCTGGGAGAAGTTCGCGCGATACTTCGAGGTTGAGCTCAAGGAAGTGAAGCTGAGTGAC 428

y          186 GlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIleCys 205
               |||||||||||||||||||||   :::|||||||||||||||||||||||||||||||||
b          429 GGCTACTACGTCATGGACCCAGCTAAGGCCGTGGATATGGTCGACGAGAACACCATCTGC 488

y          206 ValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspValLysLeuLeuAsn 225
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
b          489 GTCGCGGCGATCCTCGGGTCGACGCTGAACGGGGAGTTCGAGGACGTGAAGCTGCTCAAC 548

y          226 AspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIleHisValAspAla 245
               |||||||||   :::||||||   ||||||||||||||||||||||||||||||||||||
b          549 GATCTGCTCACCAAGAAGAACGCTGAAACAGGCTGGGACACGCCGATCCACGTGGACGCG 608

y          246 AlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGluTrpAspPheArgLeu 265
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          609 GCGAGCGGCGGGTTCATCGCGCCGTTCCTGTACCCGGAGCTGGAGTGGGACTTTCGGCTG 668

y          266 ProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeuValTyrAlaGlyIle 285
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          669 CCGCTGGTGAAGAGCATCAACGTGAGCGGGCACAAGTACGGGCTCGTCTACGCCGGGATT 728

y          286 GlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeuIlePheHisIleAsn 305
                  |||   ||||||   :::||||||||||||||||||||||||||||||||||||
b          729 CGGTGGTGCATCTGGGAGAGCAAGGAGGATCTGCCTGAGGAGCTCATCTTTCACATAACT 788

y          306 TyrLeuGlyAlaAspGlnProThrPhe 314
               |||   |||||||||:::|||   :::
b          789 TACT--GGNGCCGACAGGCCACTTTAC 813

ESULT 4
F275438
OCUS       BF275438                837 bp    mRNA    linear   EST 07-MAR-2001
EFINITION  GA__Eb0024A19f Gossypium arboreum 7-10 dpa fiber library Gossypium
           arboreum cDNA clone GA__Eb0024A19f, mRNA sequence.
CCESSION   BF275438
ERSION     BF275438.1  GI:11206508
EYWORDS    EST.
OURCE      Gossypium arboreum
 ORGANISM  Gossypium arboreum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
           ; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
EFERENCE   1  (bases 1 to 837)
 AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
           ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
 TITLE     An integrated analysis of the genetics, development, and evolution
           of the cotton fiber
 JOURNAL   Unpublished
OMMENT     Contact: Wing RA
           Clemson University Genomics Institute
           Clemson University
           100 Jordan Hall, Clemson, SC 29634, USA
           Tel: 864 656 7288
           Fax: 864 656 4293
           Email: rwing@clemson.edu
           Seq primer: TAATACGACTCACTATAGGG
           High quality sequence start: 2
           High quality sequence stop: 718.
```

```
FEATURES             Location/Qualifiers
     source          1. .837
                     /organism="Gossypium arboreum"
                     /mol_type="mRNA"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Eb0024A19f"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      244 a    171 c    198 g    223 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              3.98e-138      Length:          837
Score:                  1279.00        Matches:         248
Percent Similarity:     94.80%         Conservative:    7
Best Local Similarity:  92.19%         Mismatches:      14
Query Match:            48.91%         Indels:          1
DB:                     10             Gaps:            0

US-10-006-852-2 (1-502) x BF275438 (1-837)

Qy           1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
               ||||||||||||         ||||||:::||||||||||||||||||||||||||||||
Db          32 ATGGTGTTATCTAAGACAGCATCTGAAACTGATGTCTCAGTTCACTCCACTTTTGCCTCT 91

Qy          21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          92 CGTTATGTCCGAACCTCACTGCCCAGGTTCAAAATGCCAGAAAACTCCATACCAAAAGAG 151

Qy          41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         152 GCTGCTTATCAGATCATCAATGATGAACTGATGCTTGATGGCAACCCAAGGTTGAACCTT 211

Qy          61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
               |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db         212 GCCTCTTTTGTTACTACATGGATGGAGCCTGAATGTGATAAGCTTATAATGGACTCCATC 271

Qy          81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         272 AACAAGAACTACGTTGACATGGATGAATACCCTGTCACCACTGAGCTTCAGAATCGTTGT 331

Qy         101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
               ||||||||||||||||||||||||||||||||||||   |||:::|||   |||||||||
Db         332 GTGAACATGATAGCACATCTTTTCAACGCACCATTGGGAGAGTCAGAGGCAGCAGTTGGT 391

Qy         121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         392 GTGGGAACAGTTGGGTCATCGGAGGCAATCATGTTGGCTGGCCTTGCTTTCAAGAGAAAA 451

Qy         141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
               |||||||||||||||||||||||||||||||||   :::|||||||||||||||||||||
Db         452 TGGCAAAACAAGCGTAAGGCTGAAGGCAAACCTTATGAAAAGCCGAATATCGTGACTGGA 511

Qy         161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         512 GCCAATGTTCAGGTTTGCTGGGAGAAATTTGCAAGGTACTTTGAAGTGGAGTTGAAAGAA 571

Qy         181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
               |||||||||   ||||||||||||||||||||||||   :::||||||:::|||||||||
Db         572 GTGAAGCTTAGAGAAGGGTACTATGTGATGGACCCTGTCAAAGCAGTGGAAATGGTTGAT 631

Qy         201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         632 GAAAACACCATCTGTGTAGCTGCAATCTTGGGTTCAACCCTCAATGGAGAATTTGAAGAT 691

Qy         221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
               ||||||||||||||||||||||||:::||||||      |||||||||||||||||||||
Db         692 GTCAAGCTTCTAAACGATCTTTTAATTGAAAAGACANA-GAAACTGGATGGGATACCCCC 750
```

```
y        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        751 ATTCATGTTGATGCAGCTAGTGGTGGATTTATTGCGCCATTTTTGTACCCAGAACTTGAG 810

y        261 TrpAspPheArgLeuProLeuValLys 269
             |||||||||:::   |||   ||||||
b        811 TGGGACTTCAAGGCTCCCCCTGTGAAG 837

ESULT 5
Q165651
OCUS       BQ165651                 819 bp    mRNA    linear   EST 25-APR-2002
EFINITION  EST611520 KVKC Medicago truncatula cDNA clone pKVKC-10G4, mRNA
           sequence.
CCESSION   BQ165651
ERSION     BQ165651.1  GI:20308272
EYWORDS    EST.
OURCE      Medicago truncatula (barrel medic)
 ORGANISM  Medicago truncatula
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
           ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
           Medicago.
EFERENCE   1  (bases 1 to 819)
 AUTHORS   VandenBosch,K., Endre,G., Silverstein,K, Town,C.D., Van Aken,S.,
           Utterback,T., Cheung,F. and Fraser,C.M.
 TITLE     The Medicago truncatula 'kiloclone' set; ESTs selected and
           re-arrayed from various libraries
 JOURNAL   Unpublished
OMMENT     Contact: VandenBosch K
           Department of Plant Biology
           University of Minnesota
           220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
           Tel: 612 624 2755
           Fax: 612 625 1738
           Email: kvandenb@cbs.umn.edu
           TIGR sequence name: MTNAJ76TK Alias Clone name:MHAM-57I2 More
           information is available at: www.medicago.org
           Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
EATURES             Location/Qualifiers
    source          1..819
                    /organism="Medicago truncatula"
                    /mol_type="mRNA"
                    /cultivar="A17"
                    /db_xref="taxon:3880"
                    /clone="pKVKC-10G4"
                    /tissue_type="mixed tissues"
                    /dev_stage="various stages"
                    /lab_host="XLOLR"
                    /clone_lib="KVKC"
                    /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                    XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
                    was directionally ligated into the Unizap XR vector from
                    Stratagene and packaged using Gigapack III Gold packaging
                    extracts. Plasmids containing cDNA inserts were excised
                    from the recombinant lambda-Zap phage using Ex-assist
                    helper phage and propagated in XLOLR cells."
ASE COUNT      245 a    158 c    188 g    228 t
RIGIN

lignment Scores:
red. No.:                1.13e-137     Length:        819
core:                    1275.00       Matches:       245
ercent Similarity:       95.44%        Conservative:  6
est Local Similarity:    93.16%        Mismatches:    11
uery Match:              48.76%        Indels:        1
B:                       13            Gaps:          0

S-10-006-852-2 (1-502) x BQ165651 (1-819)

y          2 ValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArg 21
             |||||||||         ||||||||||||||||||||||||||||||||||||||||||
b         32 GTTCTCTCAAAGACAGCCTCCGAGTCTGATGTCTCTGTCCACTCAACCTTTGCTTCTCGC 91
```

```
Qy         22 TyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAla 41
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||:::|||
Db         92 TATGTCAGAACTTCACTTCCTAGATTCAAGATGCCAGAGGAGTCTATACCAAAGGATGCA 151

Qy         42 AlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        152 GCATACCAAATAATAAACGATGAATTGATGCTTGATGGAAACCCTAGATTGAATTTGGCA 211

Qy         62 SerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIleAsn 81
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        212 TCCTTTGTGACAACATGGATGGAGCCTGAGTGTGATAAACTCATCATGGCTTCCATTAAC 271

Qy         82 LysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysVal 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 AAGAACTACGTTGACATGGATGAATACCCAGTCACCACTGAGCTACAGAATCGGTGTGTT 331

Qy        102 AsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyVal 121
              |||||||||||||||||||||||||||||||||||||||   |||   ||||||||||||
Db        332 AACATGATAGCTCATCTTTTCAATGCACCACTTGAAGAGAATGAGGCTGCAGTTGGTGTT 391

Qy        122 GlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrp 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 GGCACTGTTGGCTCATCAGAGGCTATAATGTTAGCTGGATTGGCATTCAAAAGGAAGTGG 451

Qy        142 GlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAla 161
              |||||||||||||||   ||||||||||||   |||||||||||||||||||||||||||
Db        452 CAGAACAAAAGAAAACAAGAGGGAAAGCCTTATGACAAACCTAACATTGTCACTGGAGCC 511

Qy        162 AsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluVal 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 AATGTCCAGGTTTGTTGGGAGAAATTTGCAAGGTACTTTGAAGTGGAATTGAAAGAGGTG 571

Qy        182 LysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGlu 201
              |||||||||||||||||||||||||||||||||   :::||||||::::::|||||||||
Db        572 AAGCTAAGTGAAGGATACTATGTAATGGACCCTGCAAAAGCTGTGGAATTGGTAGATGAG 631

Qy        202 AsnThr-IleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspVa 221
              |||||| ||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        632 AACACCCATTTGTGTTGCTGCTATCCTTGGTTCCACACTAAATGGAGAGTTCGAAGATGT 691

Qy        221 lLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIl 241
              ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        692 TAAACGCTTAAACGATCTCCTTGTTGAAAAGAATAAGGAAACTGGATGGGACACTCCTAT 751

Qy        241 eHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGluTr 261
              ||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db        752 TCATGTGGATGCAGCTAGTGGTGGCTTCATTGCTCCATTTATTTATCCAGAACTTGAGTG 811

Qy        261 pAspPhe 263
              |||||||
Db        812 GGATTTC 818

RESULT 6
BU026594
LOCUS       BU026594                 775 bp    mRNA    linear   EST 23-AUG-2002
DEFINITION  QHG17F24.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
            clone QHG17F24, mRNA sequence.
ACCESSION   BU026594
VERSION     BU026594.1  GI:22462114
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
  ORGANISM  Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE   1  (bases 1 to 775)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
```

```
           http://compgenomics.ucdavis.edu/
 JOURNAL   Unpublished
OMMENT     Contact: Alexander Kozik [R.W.Michelmore]
           Department of Vegetable Crops, R.W.Michelmore Lab
           University of California at Davis (UCD)
           Asmundson Hall, UCD, Davis, CA 95616, USA
           Tel: 1-(530)-742-1742
           Fax: 1-(530)-752-9659
           Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
           belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
           for details.
           Plate: QHG17  row: F  column: 24.
EATURES              Location/Qualifiers
   source            1. .775
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHG17F24"
                     /lab_host="E.coli"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QH_EFGHJ sunflower RHA280
                     TAG_TISSUE=hulls
                     TAG_SEQ=GCTAGTCGGG"
ASE COUNT      224 a    140 c    191 g    219 t       1 others
RIGIN

lignment Scores:
red. No.:               6.96e-132      Length:          775
core:                   1225.00        Matches:         228
ercent Similarity:      94.57%         Conservative:    16
est Local Similarity:   88.37%         Mismatches:      14
uery Match:             46.85%         Indels:          0
B:                      13             Gaps:            0

S-10-006-852-2 (1-502) x BU026594 (1-775)

y        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

y        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
             ||||||||||||   ||||||   |||||||||   ||||||||||||||||||||||||
b         61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

y        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180

y        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
             |||||||||   ||||||||||||||||||||||||::::::||||||:::|||||||||
b        181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

y        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
             |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b        241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

y        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
             ||||||   |||||||||||||||:::|||||||||   :::||||||||||||||||||
b        301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

y        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
b        361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420
```

```
Qy       261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

Qy       281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             |||||||||||||||||||||||||||||||||||||||:::||||||||||||:::||||||
Db       481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

Qy       301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600

Qy       321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||:::||||||||||||||||||   ||||||||||||   |||||||||
Db       601 GGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

Qy       341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             :::::::::||||||||||||   ||||||      ||||||::::::||||||||||||
Db       661 AAGAGTATCATGGAAAACTGTNCAGAAAATGCAAATGTTCTAAAGAAAGGTTTGGAGAAA 720

Qy       361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPhe 378
             |||   ||||||||||||:::||||||:::   |||||||||||||||||||||
Db       721 ACCGGACGGTTCAACATCATCTCTAAAGAACACGGCGTCCCACTTGTAGCCTTC 774

RESULT 7
CA813481
LOCUS       CA813481                 786 bp    mRNA    linear   EST 11-APR-2003
DEFINITION  CA48LN10IF-E8 Cabernet Sauvignon Leaf - CA48LN Vitis vinifera cDNA
            clone CA48LN10IF-E8 5', mRNA sequence.
ACCESSION   CA813481
VERSION     CA813481.1  GI:26262418
KEYWORDS    EST.
SOURCE      Vitis vinifera
  ORGANISM  Vitis vinifera
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; Vitaceae; Vitis.
REFERENCE   1  (bases 1 to 786)
  AUTHORS   Goes da Silva,F., Lim,H., Iandolino,A., Baek,J., Jones,K., Walker
            ,M.A. and Cook,D.R.
  TITLE     Transcriptional responses of Vitis vinifera to infection by the
            bacterial pathogen Xylella fastidiosa
  JOURNAL   Unpublished
COMMENT     Contact: Doug Cook
            CAES Genome Facility
            UC Davis Department of Plant Pathology
            1 Shields Ave., DAvis, CA 95616, USA
            Tel: 530 754 6561
            Fax: 530 754 6617
            Email: drcook@ucdavis.edu
            Seq primer: GTTATCAGTCGACGGTACC.
FEATURES             Location/Qualifiers
     source          1. .786
                     /organism="Vitis vinifera"
                     /mol_type="mRNA"
                     /cultivar="Cabernet Sauvignon"
                     /db_xref="taxon:29760"
                     /clone="CA48LN10IF-E8"
                     /sex="hermaphrodite"
                     /dev_stage="Late season sample"
                     /lab_host="DH5alpha"
                     /clone_lib="Cabernet Sauvignon Leaf - CA48LN"
                     /note="Organ: Leaf; Vector: pDNR; Site_1: SfiI; Site_2:
                     SfiI; CA48LN is a cDNA library of Cabernet Sauvignon
                     leaves. The leaves were collected on September 20, 2001,
                     in Napa Valley, California, and represent leaves in late
                     season development. These leaves were asymptomatic and
                     verified to be non-infected with the bacterial pathogen,
                     Xylella fastidiosa, based on a diagnostic assay using PCR
                     and Xylella-specific primer pairs. cDNAs were made by
                     oligo-dT priming and directionally cloned. 5'and 3'
                     adaptors were used in cloning as follows:
```

```
              5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
              5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Library was
              constructed using the Clontech Creator SMART kit and
              size-selected to contain the 0.5-3 kb size fraction."
ASE COUNT       222 a    163 c    200 g    201 t
RIGIN

lignment Scores:
red. No.:              2.72e-131      Length:         786
core:                  1220.00        Matches:        232
ercent Similarity:     95.42%         Conservative:   18
est Local Similarity:  88.55%         Mismatches:     12
uery Match:            46.65%         Indels:         1
B:                     14             Gaps:           0

S-10-006-852-2 (1-502) x CA813481 (1-786)

y        175 GluValGluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGln 194
             |||||||||||||||||||||||||||::::::|||||||||||||||||||||   :::
b          1 GAGGTGGAATTGAAGGAAGTGAAGCTGACGGATGGCTACTATGTGATGGACCCTGTCAAA 60

y        195 AlaValAspMetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeu 214
             ||||||:::||||||||||||||||||||||||||||||   ||||||||||||||||||
b         61 GCAGTGGAGATGGTGGACGAGAACACCATCTGTGTTGCTGCTATCCTTGGTTCTACCCTC 120

y        215 AsnGlyGluPheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGlu 234
             ||||||||||||||||||||||||   |||||||||||||||   :::||||||||||||
b        121 AATGGAGAATTCGAAGATGTTAAACACTTGAATGATCTTTTGACAAAGAAAAACAAAGAA 180

y        235 ThrGlyTrpAspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPhe 254
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        181 ACAGGATGGGATACTCCAATCCATGTAGATGCAGCAAGTGGTGGCTTTATTGCACCCTTT 240

y        255 LeuTyrProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSer 274
             |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        241 CTTCACCCAGAGCTGGAGTGGGACTTCCGCTTGCCACTTGTGAAAAGCATTAACGTTAGC 300

y        275 GlyHisLysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGlu 294
             |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
b        301 GGCCACAAGTACGGGCTTGTGTATGCTGGCATAGGTTGGGTGATTTGGAGGAGTAAAGAG 360

y        295 AspLeuProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPhe 314
             |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
b        361 GACTTGCCCGATGAGCTCATCTTTCATATCAACTATCTTGGAGCCGACCAACCTACCTTC 420

y        315 ThrLeuAsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArg 334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        421 ACCCTCAACTTCTCCAAAGGTTCTAGTCAAGTTATTGCTCAGTATTATCAACTCATTCGC 480

y        335 LeuGlyHisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeu 354
             ||||||||||||||||||||||||:::||||||||||||:::||||||      ||||||
b        481 TTGGGCCATGAGGGATATAGGAACATTATGGAAAACTGTCAAGAAAATGCAAGGGTGCTC 540

y        355 ArgGluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValPro 374
             :::||||||||||||||||||      ||||||||||||||||||:::   |||||||||
b        541 AAAGAAGGATTGGAGAAGACAGGGCACTTCAACATTGTGTCTAAAGAGAATGGAGTGCCA 600

y        375 LeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMet 394
             ||||||||||||||||||||||||:::|||||||||   |||||||||:::||||||:::
)b       601 TTAGTAGCCTTCTCTTTGAAAGACAACAGCTGTCACAATGAGTTCGAGGTGTCAGACATA 660

)y       395 LeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIle 414
             ||||||||||||||||||||||||||||||||||||:::||||||:::|||||||||:::
)b       661 CTGCGACGCTATGGATGGATCGTGCCAGCTTATACCATACCGCCGGATGCACAACATGTG 720

)y       415 ThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuVal 434
             |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
)b       721 ACTGTTCTTCGTGTCGTCATAAGAGAAGA-TTCTCGCGCACACTTGCTGAGCGTCTGGTC 779

)y       435 IleAsp 436
                |||
)b       780 ACAGAT 785
```

```
RESULT 8
BF263784
LOCUS       BF263784                 823 bp    mRNA    linear   EST 23-OCT-2001
DEFINITION  HV_CEa0007K14f Hordeum vulgare seedling green leaf EST library
            HVcDNA0004 (Blumeria challenged) Hordeum vulgare subsp. vulgare
            cDNA clone HV_CEa0007K14f, mRNA sequence.
ACCESSION   BF263784
VERSION     BF263784.2  GI:13261070
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 823)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Wei,F., Begum,D.,
            Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi
            ,D.W., Fenton,R.D., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Blumeria infected incompatible (Mla13)
            seedling leaf cDNA library
  JOURNAL   Unpublished
COMMENT     On Nov 17, 2000 this sequence version replaced gi:11194778.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 571
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 721.
FEATURES             Location/Qualifiers
     source          1..823
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="CI16155 (Mla13)"
                     /db_xref="taxon:112509"
                     /clone="HV_CEa0007K14f"
                     /tissue_type="seedling green leaf"
                     /lab_host="TJC121"
                     /clone_lib="Hordeum vulgare seedling green leaf EST
                     library HVcDNA0004 (Blumeria challenged)"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1;
                     C.I. 16155 (Mla13) plants were greenhouse grown in the R
                     Wise lab at Iowa State University, Ames, IA; 7 day old
                     green seedlings were challenged with isolate A27 (AvrMla13
                     ) of Blumeria graminis f. sp. hordei, and leaves were
                     harvested 20 and 24 hr post-inoculation and snap frozen;
                     uninoculated leaves were harvested 20 hr post-inoculation
                     (Wei, Wise). In the TJ Close lab at the University of
                     California, Riverside, total RNA was prepared from each
                     sample pool, equal quantities of all three RNA pools were
                     combined, poly(A) RNA was purified from the mixture, one
                     cDNA library was made, and 1 million pfu were in vivo
                     excised to give pBluescript SK(-) cDNA phagemids (Choi,
                     Close). Phagemids were plated and picked at the Clemson
                     University Genomics Institute (CUGI) (Begum, Palmer,
                     Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
                     sequencing and sequence analysis were performed at CUGI
                     (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
                     The sequence has been trimmed to remove vector sequence
                     and contains a minimum of 100 bases of phred value 20 or
                     above. For more details on library preparation and
                     sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT      209 a    187 c    249 g    175 t      3 others
```

```
RIGIN

lignment Scores:
red. No.:               5.71e-131      Length:         823
core:                   1217.50        Matches:        236
ercent Similarity:      90.91%         Conservative:   14
est Local Similarity:   85.82%         Mismatches:     21
uery Match:             46.56%         Indels:         4
B:                      10             Gaps:           1

S-10-006-852-2 (1-502) x BF263784 (1-823)

y         23 ValArgThrSerLeuProArgPhe---LysMetPro--GluAsnSerIleProLysGluA 41
             ||||||      |||||||||         ||||||   |||||||||||||||||||||
b          5 GTGCGCGAGGAGCTCCCGCGGTANCCGGNNATGCCGGAGAAAACTCGATCCCCAAGGAGG 64

y         41 laAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuA 61
             |||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
b         65 CGGCGTACCAGATCATCAGCGACGAGCTGATGCTGGACGGCAACCCGCGGCTGAACCTGG 124

y         61 laSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIleA 81
             |||||||||||||||||||||||||||||||||||    |||||||||||   |||:::|
b        125 CGTCCTTCGTCACCACCTGGATGGAGCCCGAGTGCGGCAAGCTCATCATGGACTCCGTCA 184

y         81 snLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysV 101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        185 ACAAGAACTACGTCGACATGGACGAGTACCCCGTCACCACCGAGCTCCAGAACCGTTGCG 244

y        101 alAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyV 121
             |||||||||||||||||||||||||||||||||:::   |||   |||||||||:::|||
b        245 TAAACATGATAGCTCACTTGTTCAATGCACCGATCGGCGAGGACGAGACAGCTATCGGAG 304

y        121 alGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysT 141
             ||   |||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        305 TCTCGACGGTGGGGTCTTCGGAAGCAATCATGCTTGCAGGCCTGGCGTTCAAGAGGAAGT 364

y        141 rpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyA 161
             ||   ||||||    |||   :::|||||||||   ||||||||||||||||||||||||
b        365 GGGCGAACAAAATGAAGGAGCAGGGGAAGCCATGCGACAAACCTAACATTGTTACTGGTG 424

y        161 laAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluV 181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        425 CAAATGTTCAGGTTTGCTGGGAGAAATTTGCTAGGTATTTTGAAGTGGAATTGAAGGAGG 484

y        181 alLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspG 201
             ||||||||:::||||||||||||||||||||||||::::::||||||:::||||||||||
b        485 TCAAGTTGACTGAAGGGTACTATGTCATGGATCCTAAGAAGGCTGTTGAAATGGTGGATG 544

y        201 luAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspV 221
             ||||||||||||||||||||   ||||||||||||||||||   ||||||:::|||||||
b        545 AGAACACTATATGTGTCGCCGCCATCCTGGGATCTACTCTCACTGGAGAGTACGAAGATG 604

y        221 alLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProI 241
             ||||||||||||||||||||||||||||||||||||||||||||||||||:::   ||||
b        605 TCAAACTGTTGAATGACCTTCTTGTGGAGAAGAACAAAGAAACAGGGTGGAACGTGCCGA 664

y        241 leHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGluT 261
             ||||||||||||||||||||||||||||||||||||||||||||   |||||||||||||
b        665 TCCATGTTGATGCTGCCAGCGGAGGATTTATCGCTCCGTTTCTTTAGCCTGAGCTTGAAT 724

)y       261 rpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyr-GlyLeu 280
             ||||||||:::||||||||||||||||||||||||      ||||||||||||    |||
)b       725 GGGACTTCAAGCTACCATTGGTGAAGAGCATCAACGGTAGGGGGCACAAGTATTGCCCTT 784

)y       281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLys 293
             ||||||   |||:::||||||||||||||||||:::|||
)b       785 GTGTACCCTGGGGTTGGATGGGTCATTTGGCGGAGCAAG 823

RESULT 9
BQ971999
LOCUS       BQ971999                 762 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  QHB9G09.yg.ab1 QH_ABCDI sunflower RHA801 Helianthus annuus cDNA
            clone QHB9G09, mRNA sequence.
```

```
ACCESSION   BQ971999
VERSION     BQ971999.1  GI:22389520
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
  ORGANISM  Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QHB9  row: G  column: 09.
FEATURES             Location/Qualifiers
     source          1..762
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA801"
                     /db_xref="taxon:4232"
                     /clone="QHB9G09"
                     /lab_host="E.coli"
                     /clone_lib="QH_ABCDI sunflower RHA801"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QH_ABCDI sunflower RHA801
                     TAG_TISSUE=hulls
                     TAG_SEQ=GCTAGTCGGG"
BASE COUNT      224 a    138 c    190 g    210 t
ORIGIN

Alignment Scores:
Pred. No.:               2.48e-129      Length:         762
Score:                   1203.00        Matches:        224
Percent Similarity:      94.09%         Conservative:   15
Best Local Similarity:   88.19%         Mismatches:     15
Query Match:             46.00%         Indels:         0
DB:                      13             Gaps:           0

US-10-006-852-2 (1-502) x BQ971999 (1-762)

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||   |||||||||   ||||||||||||||||||||||||
Db         61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180
```

```
y           181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
                |||||||||   ||||||||||||||||||||||||::::::||||||:::|||||||||
o           181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

y           201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
                |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b           241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

y           221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
                ||||||   |||||||||||||||:::|||||||||   :::||||||||||||||||||
b           301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

y           241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
b           361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

y           261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b           421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

y           281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
                |||||||||||||||||||||||||||||||||||||||:::|||||||||:::||||||
b           481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

y           301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b           541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600

y           321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
                ||||||||||||   ||||||||||||||||||   ||||||||||||   |||||||||
b           601 GGTTCTAGTCAAACAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

y           341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
                :::::::::||||||||||||:::||||||      ||||||::::::||||||||||||
b           661 AAGAGCATCATGGAAAACTGTCAAGAAAATGCAAATGTTCTAAAGAAAGGTTTGGAGAAA 720

y           361 ThrGluArgPheAsnIleValSerLysAspGluGlyValPro 374
                |||   |||   ||||||:::|||||||||   |||||||||
b           721 ACCGGACGGGTCAACATCATCTCTAAAGACAACGGCGTCCCA 762

ESULT 10
E054156
OCUS       BE054156                803 bp    mRNA    linear   EST 07-MAR-2001
EFINITION  GA__Ea0034H22f Gossypium arboreum 7-10 dpa fiber library Gossypium
           arboreum cDNA clone GA__Ea0034H22f, mRNA sequence.
CCESSION   BE054156
ERSION     BE054156.2  GI:13246539
EYWORDS    EST.
OURCE      Gossypium arboreum
 ORGANISM  Gossypium arboreum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
           ; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
EFERENCE   1  (bases 1 to 803)
 AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
           ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
 TITLE     An integrated analysis of the genetics, development, and evolution
           of the cotton fiber
 JOURNAL   Unpublished
OMMENT     On Jun 8, 2000 this sequence version replaced gi:8381212.
           Contact: Wing RA
           Clemson University Genomics Institute
           Clemson University
           100 Jordan Hall, Clemson, SC 29634, USA
           Tel: 864 656 7288
           Fax: 864 656 4293
           Email: rwing@clemson.edu
           Seq primer: TAATACGACTCACTATAGGG
           High quality sequence stop: 792.
EATURES             Location/Qualifiers
     source          1..803
                     /organism="Gossypium arboreum"
                     /mol_type="mRNA"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Ea0034H22f"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      236 a    152 c    215 g    198 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              1.76e-128       Length:         803
Score:                  1196.00         Matches:        220
Percent Similarity:     92.86%          Conservative:   27
Best Local Similarity:  82.71%          Mismatches:     19
Query Match:            45.74%          Indels:         0
DB:                     10              Gaps:           0

US-10-006-852-2 (1-502) x BE054156 (1-803)

Qy          122 GlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrp 141
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            2 GGGACAGTTGGTTCCTCTGAGGCAATAATGCTGGCAGGATTAGCCTTTAAAAGGAAGTGG 61

Qy          142 GlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAla 161
                |||   |||   |||::::::|||||||||   ||||||||||||||||||||||||||||||
Db           62 CAACAAAAGATGAAATCACAGGGAAAACCGTATGATAAGCCTAACATAGTCACCGGAGCG 121

Qy          162 AsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluVal 181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          122 AATGTGCAGGTTTGTTGGGAGAAGTTTGCGAGGTATTTCGAGGTTGAATTGAAGGAAGTG 181

Qy          182 LysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGlu 201
                ||||||   ||||||||||||||||||||||||   :::||||||:::||||||||||||
Db          182 AAGTTGAAAGAGGGATACTATGTGATGGACCCTGTGAAAGCGGTTGAAATGGTTGACGAG 241

Qy          202 AsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspVal 221
                ||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::|||
Db          242 AATACCATATGTGTTGCAGCAATTCTCGGATCCACCCTAACAGGGGAGTTTGAGAATGTG 301

Qy          222 LysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle 241
                ||||||||||||:::||||||   :::|||||||||||||||||||||||||||||||||
Db          302 AAGCTCCTTAATGAACTTCTCACCAAGAAGAATAAGGAAACCGGTTGGGATACCCCAATA 361

Qy          242 HisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGluTrp 261
                ||||||||||||||||||||||||||||||||||||||||||:::||||||::::::||||||
Db          362 CATGTGGATGCTGCTAGTGGAGGGTTCATTGCTCCCTTTGTTTACCCGAAAATCGAATGG 421

Qy          262 AspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeuVal 281
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          422 GATTTCCGCCTGCCGTTAGTAAAAAGCATCAATGTCAGTGGGCACAAGTATGGCCTTGTA 481

Qy          282 TyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeuIle 301
                |||||||||||||||||||||:::||||||:::||||||||||||||||||::::::|||:::
Db          482 TATGCTGGTATCGGTTGGGTCGTGTGGAGAAGCAAGGAGGATTTGCCAGACGATCTCGTA 541

Qy          302 PheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLysGly 321
                |||||||||||||||||||||:::|||||||||||||||||||||   ||||||||||||
Db          542 TTTCACATCAACTACCTTGGATCCGATCAGCCCACATTCACCTTANACTTCTCTAAAGGC 601

Qy          322 SerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyrArg 341
                |||||||||:::|||||||||||||||||||||||||||||||||   |||||||||:::
Db          602 TCTAGTCAAATCATAGCTCAATATTATCAGCTTATTCGGCTCGGTTNTGAGGGATACAAG 661

Qy          342 AsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLysThr 361
                ::::::||||||||||||      |||      :::|||:::||||||:::|||||||||
Db          662 AGTATCATGGAAAACTGCATCGGGAATGCAAGAATCCTGAAGGAAGGGATAGAAAAAACG 721

Qy          362 GluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeuLys 381
                   ||||||   :::||||||||||||   ||||||||||||   ||||||:::||||||
Db          722 GGACGATTCGAAGTCGTCTCCAAAGACGTGGGCGTTCCCCTCGGGGCTTTTGCTCTAAAA 781
```

```
y         382 AspSerSerCysHisThr 387
              ||||||:::   :::|||
b         782 GACAGCACTAAATACACG 799

ESULT 11
U023643
OCUS       BU023643                  750 bp    mRNA    linear   EST 23-AUG-2002
EFINITION  QHF12C03.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
           clone QHF12C03, mRNA sequence.
CCESSION   BU023643
ERSION     BU023643.1  GI:22459163
EYWORDS    EST.
OURCE      Helianthus annuus (common sunflower)
 ORGANISM  Helianthus annuus
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           asterids; campanulids; Asterales; Asteraceae; Asteroideae;
           Heliantheae; Helianthus.
EFERENCE   1  (bases 1 to 750)
 AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
           Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
           ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
           Church,S., Jackson,L. and Bradford,K.
 TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
           http://compgenomics.ucdavis.edu/
 JOURNAL   Unpublished
OMMENT     Contact: Alexander Kozik [R.W.Michelmore]
           Department of Vegetable Crops, R.W.Michelmore Lab
           University of California at Davis (UCD)
           Asmundson Hall, UCD, Davis, CA 95616, USA
           Tel: 1-(530)-742-1742
           Fax: 1-(530)-752-9659
           Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
           belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
           for details.
           Plate: QHF12  row: C  column: 03.
EATURES             Location/Qualifiers
    source          1..750
                    /organism="Helianthus annuus"
                    /mol_type="mRNA"
                    /cultivar="RHA280"
                    /db_xref="taxon:4232"
                    /clone="QHF12C03"
                    /lab_host="E.coli"
                    /clone_lib="QH_EFGHJ sunflower RHA280"
                    /note="Vector: pBRcDNASfiAB; The library was constructed
                    from 11 different sources of RNA from a single genotype.
                    Separate cDNAs were generated using primers that
                    incorporated unique 5' and 3' tags to distinguish each
                    source of RNA. cDNAs were then pooled, size-fractionated,
                    directionally cloned into a custom medium-copy vector and
                    transformations made with four size classes to minimize
                    size bias. Details of each source of RNA and library
                    construction can be obtained at http://cgpdb.ucdavis.edu/
                    TAG_SEQ=Not found"
ASE COUNT      221 a    132 c    187 g    210 t
RIGIN

lignment Scores:
red. No.:              1.04e-127      Length:        750
core:                  1189.00        Matches:       221
ercent Similarity:     94.80%         Conservative:  16
est Local Similarity:  88.40%         Mismatches:    13
uery Match:            45.47%         Indels:        0
B:                     13             Gaps:          0

S-10-006-852-2 (1-502) x BU023643 (1-750)

y         121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b           1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60
```

```
Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||   |||||||||   ||||||||||||||||||||||||
Db         61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              |||||||||   ||||||||||||||||||||||||::::::||||||:::|||||||||
Db        181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   |||||||||||||||:::|||||||||   :::||||||||||||||||||
Db        301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              |||||||||||||||||||||||||||||||||||||||:::|||||||||:::||||||
Db        481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600

Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||:::||||||||||||||||||   ||||||||||||   |||||||||
Db        601 GGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

Qy        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              :::::::::||||||||||||:::||||||      ||||||::::::||||||||||||
Db        661 AAGAGCATCATGGAAAACTGTCAAGAAAATGCAAATGTTCTAAAGAAAGGTTTGGAGAAA 720

Qy        361 ThrGluArgPheAsnIleValSerLysAsp 370
              |||   |||   ||||||:::|||||||||
Db        721 ACCGGACGGGTCAACATCATCTCTAAAGAC 750

RESULT 12
BU027739
LOCUS       BU027739                  786 bp    mRNA    linear   EST 23-AUG-2002
DEFINITION  QHG7H16.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
            clone QHG7H16, mRNA sequence.
ACCESSION   BU027739
VERSION     BU027739.1  GI:22463259
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
  ORGANISM  Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE   1  (bases 1 to 786)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
```

```
           Asmundson Hall, UCD, Davis, CA 95616, USA
           Tel: 1-(530)-742-1742
           Fax: 1-(530)-752-9659
           Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
           belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
           for details.
           Plate: QHG7  row: H  column: 16.
EATURES             Location/Qualifiers
    source          1. .786
                    /organism="Helianthus annuus"
                    /mol_type="mRNA"
                    /cultivar="RHA280"
                    /db_xref="taxon:4232"
                    /clone="QHG7H16"
                    /lab_host="E.coli"
                    /clone_lib="QH_EFGHJ sunflower RHA280"
                    /note="Vector: pBRcDNASfiAB; The library was constructed
                    from 11 different sources of RNA from a single genotype.
                    Separate cDNAs were generated using primers that
                    incorporated unique 5' and 3' tags to distinguish each
                    source of RNA. cDNAs were then pooled, size-fractionated,
                    directionally cloned into a custom medium-copy vector and
                    transformations made with four size classes to minimize
                    size bias. Details of each source of RNA and library
                    construction can be obtained at http://cgpdb.ucdavis.edu/
                    TAG_LIB=QH_EFGHJ sunflower RHA280
                    TAG_TISSUE=hulls
                    TAG_SEQ=GCTAGTCGGG"
ASE COUNT      228 a    142 c    194 g    220 t      2 others
RIGIN

lignment Scores:
red. No.:              1.67e-127      Length:         786
core:                  1187.50        Matches:        227
ercent Similarity:     92.37%         Conservative:   15
est Local Similarity:  86.64%         Mismatches:     18
uery Match:            45.41%         Indels:         2
B:                     13             Gaps:           1

S-10-006-852-2 (1-502) x BU027739 (1-786)

y        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

y        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
             ||||||||||||   ||||||   |||||||||   ||||||||||||||||||||||||
b         61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

y        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180

y        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
             |||||||||   ||||||||||||||||||||||||::::::||||||:::|||||||||
b        181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

y        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
             |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b        241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

Qy       221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
             ||||||   |||||||||||||||:::|||||||||   :::||||||||||||||||||
Db       301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

Qy       241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db       361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

Qy       261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480
```

```
Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              |||||||||||||||||||||||||||||||||||||||:::|||||||||:::||||||
Db        481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600

Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeu-GlyHisGluGlyTy 340
              ||||||||||||:::||||||||||||||||||   |||||||||       ||||||||
Db        601 GGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGNGTTTCGAGGGATA 660

Qy        340 rArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeu---Gl 359
              |   ::::::||||||||||||:::||||||      ||||||:::|||   :::     
Db        661 CNAGAGCATCATGGAAAACTGTCAAGAAAATGCAAATGTTCTAAAAGAAAGGGTTTGGAG 720

Qy        359 uLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSe 379
               |||      ||||||||||||:::|||||||||   |||||||||||||||||||||||
Db        721 AAAACCGGGACGGTTCAACATCATCTCTAAAGACAACGGCGTCCCACTTGTAGCCTTCTC 780

Qy        379 rLeu 380
              ||||
Db        781 GTTA 784

RESULT 13
BU025264
LOCUS       BU025264                 746 bp    mRNA    linear   EST 23-AUG-2002
DEFINITION  QHF8J20.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
            clone QHF8J20, mRNA sequence.
ACCESSION   BU025264
VERSION     BU025264.1  GI:22460784
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
  ORGANISM  Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE   1  (bases 1 to 746)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QHF8  row: J  column: 20.
FEATURES             Location/Qualifiers
     source          1. .746
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHF8J20"
                     /lab_host="E.coli"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
```

```
                 construction can be obtained at http://cgpdb.ucdavis.edu/
                 TAG_LIB=QH_EFGHJ sunflower RHA280
                 TAG_TISSUE=hulls
                 TAG_SEQ=GCTAGTCGGG"
ASE COUNT      219 a    131 c    185 g    211 t
RIGIN

lignment Scores:
red. No.:               3.01e-127      Length:        746
core:                   1185.00        Matches:       220
ercent Similarity:      95.16%         Conservative:  16
est Local Similarity:   88.71%         Mismatches:    12
uery Match:             45.32%         Indels:        0
B:                      13             Gaps:          0

S-10-006-852-2 (1-502) x BU025264 (1-746)

y        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

y        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
             ||||||||||||   ||||||   |||||||||   ||||||||||||||||||||||||
b         61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

y        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180

y        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
             |||||||||   ||||||||||||||||||||||||::::::||||||:::|||||||||
b        181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

y        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
             |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b        241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

y        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
             ||||||   |||||||||||||||:::|||||||||   :::||||||||||||||||||
b        301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

y        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
b        361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

y        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

y        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             |||||||||||||||||||||||||||||||||||||||:::|||||||||:::||||||
b        481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

y        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600

Qy       321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||:::||||||||||||||||||   ||||||||||||   |||||||||
Db       601 GGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

Qy       341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             :::::::::||||||||||||:::||||||      ||||||::::::||||||||||||
Db       661 AAGAGCATCATGGAAAACTGTCAAGAAAATGCAAATGTTCTAAAGAAAGGTTTGGAGAAA 720

Qy       361 ThrGluArgPheAsnIleValSer 368
             |||   ||||||||||||:::|||
Db       721 ACCGGACGGTTCAACATCATCTCT 744

RESULT 14
BU027680
LOCUS       BU027680                746 bp    mRNA    linear   EST 23-AUG-2002
DEFINITION  QHG7B13.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
            clone QHG7B13, mRNA sequence.
ACCESSION   BU027680
VERSION     BU027680.1  GI:22463200
KEYWORDS    EST.
SOURCE      Helianthus annuus (common sunflower)
  ORGANISM  Helianthus annuus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Asterales; Asteraceae; Asteroideae;
            Heliantheae; Helianthus.
REFERENCE   1  (bases 1 to 746)
  AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
            Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
            ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
            Church,S., Jackson,L. and Bradford,K.
  TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
            http://compgenomics.ucdavis.edu/
  JOURNAL   Unpublished
COMMENT     Contact: Alexander Kozik [R.W.Michelmore]
            Department of Vegetable Crops, R.W.Michelmore Lab
            University of California at Davis (UCD)
            Asmundson Hall, UCD, Davis, CA 95616, USA
            Tel: 1-(530)-742-1742
            Fax: 1-(530)-752-9659
            Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
            belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
            for details.
            Plate: QHG7  row: B  column: 13.
FEATURES             Location/Qualifiers
     source          1..746
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHG7B13"
                     /lab_host="E.coli"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QH_EFGHJ sunflower RHA280
                     TAG_TISSUE=hulls
                     TAG_SEQ=GCTAGTCGGG"
BASE COUNT      219 a    131 c    185 g    211 t
ORIGIN

Alignment Scores:
Pred. No.:              3.01e-127      Length:        746
Score:                  1185.00        Matches:       220
Percent Similarity:     95.16%         Conservative:  16
Best Local Similarity:  88.71%         Mismatches:    12
Query Match:            45.32%         Indels:        0
DB:                     13             Gaps:          0

US-10-006-852-2 (1-502) x BU027680 (1-746)

Qy       121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

Qy       141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
             ||||||||||||   ||||||   |||||||||   ||||||||||||||||||||||||
Db        61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

Qy       161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180
```

```
y         181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              |||||||||   ||||||||||||||||||||||||::::::||||||:::|||||||||
b         181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

y         201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b         241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

y         221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   |||||||||||||||:::|||||||||   :::||||||||||||||||||
b         301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

y         241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
b         361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

y         261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

y         281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              |||||||||||||||||||||||||||||||||||||||:::|||||||||:::||||||
b         481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

y         301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600

y         321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||:::||||||||||||||||||   ||||||||||||   |||||||||
b         601 GGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

y         341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              :::::::::||||||||||||:::||||||      ||||||::::::||||||||||||
b         661 AAGAGCATCATGGAAAACTGTCAAGAAAATGCAAATGTTCTAAAGAAAGGTTTGGAGAAA 720

y         361 ThrGluArgPheAsnIleValSer 368
              |||   ||||||||||||:::|||
b         721 ACCGGACGGTTCAACATCATCTCT 744

ESULT 15
U025649
OCUS       BU025649                 746 bp    mRNA    linear   EST 23-AUG-2002
EFINITION  QHG10J06.yg.ab1 QH_EFGHJ sunflower RHA280 Helianthus annuus cDNA
           clone QHG10J06, mRNA sequence.
CCESSION   BU025649
ERSION     BU025649.1  GI:22461169
EYWORDS    EST.
OURCE      Helianthus annuus (common sunflower)
 ORGANISM  Helianthus annuus
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           asterids; campanulids; Asterales; Asteraceae; Asteroideae;
           Heliantheae; Helianthus.
EFERENCE   1  (bases 1 to 746)
 AUTHORS   Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
           Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegle,J., Ellison
           ,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
           Church,S., Jackson,L. and Bradford,K.
 TITLE     Lettuce and Sunflower ESTs from the Compositae Genome Project
           http://compgenomics.ucdavis.edu/
 JOURNAL   Unpublished
OMMENT     Contact: Alexander Kozik [R.W.Michelmore]
           Department of Vegetable Crops, R.W.Michelmore Lab
           University of California at Davis (UCD)
           Asmundson Hall, UCD, Davis, CA 95616, USA
           Tel: 1-(530)-742-1742
           Fax: 1-(530)-752-9659
           Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
           belongs to contig QH_CA_Contig2686, see http://cgpdb.ucdavis.edu/
           for details.
           Plate: QHG10  row: J  column: 06.
```

```
FEATURES             Location/Qualifiers
     source          1..746
                     /organism="Helianthus annuus"
                     /mol_type="mRNA"
                     /cultivar="RHA280"
                     /db_xref="taxon:4232"
                     /clone="QHG10J06"
                     /lab_host="E.coli"
                     /clone_lib="QH_EFGHJ sunflower RHA280"
                     /note="Vector: pBRcDNASfiAB; The library was constructed
                     from 11 different sources of RNA from a single genotype.
                     Separate cDNAs were generated using primers that
                     incorporated unique 5' and 3' tags to distinguish each
                     source of RNA. cDNAs were then pooled, size-fractionated,
                     directionally cloned into a custom medium-copy vector and
                     transformations made with four size classes to minimize
                     size bias. Details of each source of RNA and library
                     construction can be obtained at http://cgpdb.ucdavis.edu/
                     TAG_LIB=QH_EFGHJ sunflower RHA280
                     TAG_TISSUE=hulls
                     TAG_SEQ=GCTAGTCGGG"
BASE COUNT      218 a    132 c    185 g    211 t
ORIGIN

Alignment Scores:
Pred. No.:              3.94e-127      Length:         746
Score:                  1184.00        Matches:        220
Percent Similarity:     95.16%         Conservative:   16
Best Local Similarity:  88.71%         Mismatches:     12
Query Match:            45.28%         Indels:         0
DB:                     13             Gaps:           0

US-10-006-852-2 (1-502) x BU025649 (1-746)

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTGGAACTGTGGGCTCATCCGAAGCCATCATGTTGGCCGGACTAGCTTTCAAAAGAAAA 60

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||   |||||||||   ||||||||||||||||||||||||
Db         61 TGGCAAAACAAAATGAAAGCTCTTGGCAAACCTTGGGACAAACCTAACATTGTAACCGGG 120

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCAACGTTCAGGTTTGTTGGGAGAAATTTGCTCGGTATTTTGAAGTGGAGTTGAAGGAA 180

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              |||||||||   ||||||||||||||||||||||||::::::||||||:::|||||||||
Db        181 GTGAAGTTGAGGGAAGGTTACTACGTGATGGATCCTGAGAAAGCTGTGGAGATGGTTGAT 240

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        241 GAAAACACTATTTGTGTGGCTGCTATCTTGGGTTCCACCCTCAATGGTGAATTTGAAGAT 300

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   |||||||||||||||:::|||||||||   :::||||||||||||||||||
Db        301 GTTAAGCGTTTAAATGACCTCTTGATTGAGAAAAATGCCAAGACCGGATGGGATACACCT 360

Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        361 ATTCATGTGGACGCTGCAAGTGGAGGTTTTATTGCACCATTTATTTACCCGGAACTTGAA 420

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGGATTTCCGGTTGCCATTGGTTAAGAGTATAAATGTTAGTGGTCACAAATATGGGCTT 480

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              |||||||||||||||||||||||||||||||||||||||:::|||||||||:::||||||
Db        481 GTTTACGCCGGAATTGGGTGGGTCATTTGGAGGAACAAAGATGACTTGCCTGATGAACTC 540

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCTTTCACATCAACTATCTTGGTGCTGATCAACCAACTTTCACTCTCAACTTCTCCAAA 600
```

```
y        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||:::||||||||||||||||||   ||||||||||||   |||||||||
b        601 GGTTCTAGTCAAATAATTGCTCAATACTATCAGTTCATTCGCTTGGGTTTCGAGGGATAC 660

y        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             :::::::::||||||||||||:::||||||      ||||||::::::||||||||||||
b        661 AAGAGCATCATGGAAAACTGTCAAGAAAATGCAAATGTTCTACAGAAAGGTTTGGAGAAA 720

y        361 ThrGluArgPheAsnIleValSer 368
             |||   ||||||||||||:::|||
b        721 ACCGGACGGTTCAACATCATCTCT 744

earch completed: October 22, 2003, 17:27:39
ob time : 2452 secs
```

```
                         GenCore version 5.1.6
               Copyright (c) 1993 - 2003  Compugen Ltd.

M protein - protein search, using sw model

un on:         October 22, 2003, 14:50:13 ; Search time 69 Seconds
                                           (without alignments)
                                           1877.426 Million cell updates/sec

itle:          US-10-006-852-2
erfect score:  2615
equence:       1 MVLSHAVSESDVSVHSTFAS..........DIITGWKKFVADRKKTSGIC 502

coring table:  BLOSUM62
               Gapop 10.0 , Gapext 0.5

earched:       830525 seqs, 258052604 residues

otal number of hits satisfying chosen parameters:      830525

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :       SPTREMBL_23:*
               1:  sp_archea:*
               2:  sp_bacteria:*
               3:  sp_fungi:*
               4:  sp_human:*
               5:  sp_invertebrate:*
               6:  sp_mammal:*
               7:  sp_mhc:*
               8:  sp_organelle:*
               9:  sp_phage:*
               10:  sp_plant:*
               11:  sp_rodent:*
               12:  sp_virus:*
               13:  sp_vertebrate:*
               14:  sp_unclassified:*
               15:  sp_rvirus:*
               16:  sp_bacteriap:*
               17:  sp_archeap:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
esult          Query
 No.   Score   Match Length DB  ID                        Description
-----------------------------------------------------------------------
   1    2240   85.7    496  10  O81102                    O81102 nicotiana t
   2    2233   85.4    496  10  Q8LKR4                    Q8lkr4 nicotiana t
   3    2232   85.4    496  10  P93369                    P93369 nicotiana t
   4    2228   85.2    496  10  Q9AT17                    Q9at17 nicotiana t
   5    2199   84.1    496  10  O81101                    O81101 nicotiana t
   6  2191.5   83.8    513  10  Q8LLP2                    Q8llp2 oryza sativ
   7  2186.5   83.6    491  10  Q94KK8                    Q94kk8 nicotiana t
   8  2135.5   81.7    493  10  Q9ZPS3                    Q9zps3 arabidopsis
   9    2126   81.3    494  10  Q944L6                    Q944l6 arabidopsis
  10    2080   79.5    500  10  Q9ZPS4                    Q9zps4 arabidopsis
  11  2056.5   78.6    501  10  Q9AQU4                    Q9aqu4 oryza sativ
  12    1969   75.3    494  10  Q9LSH2                    Q9lsh2 arabidopsis
  13    1965   75.1    494  10  Q8LFR4                    Q8lfr4 arabidopsis
  14  1845.5   70.6    419  10  Q8RXH0                    Q8rxh0 arabidopsis
  15    1844   70.5    500  10  Q9AR41                    Q9ar41 oryza sativ
  16    1321   50.5    520   3  Q8X0B0                    Q8x0b0 neurospora
```

```
  17    1318   50.4    514   3  Q9UVD7                    Q9uvd7 aspergillus
  18  1251.5   47.9    467  16  P73043                    P73043 synechocyst
  19    1110   42.4    460  16  O06249                    O06249 mycobacteri
  20    1108   42.4    475  16  Q9X8J5                    Q9x8j5 streptomyce
  21    1100   42.1    468  17  Q8TPG4                    Q8tpg4 methanosarc
  22  1070.5   40.9    282  10  Q8RVS4                    Q8rvs4 oryza sativ
  23    1008   38.5    489  16  Q8FHG5                    Q8fhg5 escherichia
  24  1007.5   38.5    464  16  Q8XIQ6                    Q8xiq6 clostridium
  25     952   36.4    464   2  Q8GF15                    Q8gf15 edwardsiell
  26   635.5   24.3    304  16  Q8YBJ0                    Q8ybj0 brucella me
  27   613.5   23.5    177  10  Q9LSH6                    Q9lsh6 arabidopsis
  28     487   18.6    219   2  Q8L3L1                    Q8l3l1 lactococcus
  29     401   15.3    163   2  Q8L3L0                    Q8l3l0 lactococcus
  30     379   14.5    157   2  Q9LC52                    Q9lc52 lactococcus
  31   335.5   12.8    167  16  Q8YBI9                    Q8ybi9 brucella me
  32     325   12.4    589   3  Q05567                    Q05567 saccharomyc
  33     317   12.1    372  17  Q8TV92                    Q8tv92 methanopyru
  34     317   12.1    576   3  Q8X074                    Q8x074 neurospora
  35     314   12.0    384  17  Q9UZD5                    Q9uzd5 pyrococcus
  36     312   11.9    363  17  O27188                    O27188 methanobact
  37     311   11.9    367  17  O28275                    O28275 archaeoglob
  38     305   11.7    383  17  O58679                    O58679 pyrococcus
  39   299.5   11.5    545   5  Q9V7Y2                    Q9v7y2 drosophila
  40   290.5   11.1    395  17  Q8TUQ9                    Q8tuq9 methanosarc
  41     288   11.0    552   5  Q9Y194                    Q9y194 caenorhabdi
  42     271   10.4    398  17  Q8PXA5                    Q8pxa5 methanosarc
  43     270   10.3    544  10  Q9C509                    Q9c509 arabidopsis
  44   269.5   10.3    488  17  O28946                    O28946 archaeoglob
  45   264.5   10.1    542   5  Q17456                    Q17456 caenorhabdi


                              ALIGNMENTS


RESULT 1
O81102
ID   O81102      PRELIMINARY;      PRT;   496 AA.
AC   O81102;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Glutamate decarboxylase isozyme 1.
GN   NTGAD1.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yun S.J., Oh S.H.;
RT   "Cloning and characterization of a tobacco cDNA encoding
RT   calcium/calmodulin-dependent glutamate decarboxylase.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC       TYRDC).
DR   EMBL; AF020425; AAC24195.1; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
KW   Decarboxylase; Lyase; Pyridoxal phosphate.
SQ   SEQUENCE   496 AA;  55963 MW;  4E8492F5DF0BDA8A CRC64;

  Query Match             85.7%;  Score 2240;  DB 10;  Length 496;
  Best Local Similarity   85.1%;  Pred. No. 4.1e-161;
  Matches  428;  Conservative   35;  Mismatches   32;  Indels    8;  Gaps    3;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   ||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||:||:| |||||||||||||||||||||||||||||||||| : |||||
```

```
b          61 ASFVTTWMEPECNKLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

y         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
b         121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTGANVQVCWEKFARYFEVELKE 180

)y        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
b         181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

y         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||:||
b         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

y         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||:||||||:||  ||||||||
b         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGLEK 360

y         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              : ||||:||: |||||||||||:| | |||||: |||:|||:|||||||||||:||||||
b         361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIIPAYTMPPNAQHVTVLRVV 420

y         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||||||||: ||| ||:||  |:::    :|:|   :    ||:| :
b         421 IREDFSRTLAERLVIDIEKVLHELDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

y         481 QRDIITGWKKFVAD-RKKTSGIC 502
              | :| | ||||||| :|||:|:|
b         474 QLEITTAWKKFVADKKKKTNGVC 496

ESULT 2
8LKR4
D   Q8LKR4      PRELIMINARY;      PRT;   496 AA.
C   Q8LKR4;
T   01-OCT-2002 (TrEMBLrel. 22, Created)
T   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
T   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E   Glutamate decarboxylase.
S   Nicotiana tabacum (Common tobacco).
C   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
C   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
X   NCBI_TaxID=4097;
N   [1]
P   SEQUENCE FROM N.A.
C   STRAIN=cv. Samsun NN;
A   McLean M.D., Yevtushenko D., Deschene A., Van Cauwenberghe O.R.,
A   Makhmoudova A., Potter J.W., Bown A.W., Shelp B.J.;
T   "Transgenic tobacco plants overexpressing glutamate decarboxylase are
T   resistant to the root-knot nematode.";
L   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
C   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
C   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
C       TYRDC).
R   EMBL; AF506366; AAM48129.1; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
W   Decarboxylase; Lyase; Pyridoxal phosphate.
Q   SEQUENCE   496 AA;  55966 MW;  4771CF71BC478349 CRC64;

 Query Match            85.4%;  Score 2233;  DB 10;  Length 496;
 Best Local Similarity  84.9%;  Pred. No. 1.4e-160;
 Matches  427;  Conservative   35;  Mismatches   33;  Indels    8;  Gaps    3;

y           1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   ||||||:||||||||||||||||||||||||||||||||||||||||||||||
b           1 MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

y          61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||: |:| |||||||||||||||||||||||||||||||||| : |||||
b          61 ASFVTTWMEPECNTLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120
```

```
Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPFDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
Db        181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||:||
Db        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||:||||||:||  |||||:||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGIEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              : ||||:||: |||||||||||:| | |||||: |||:|||||||||||||||:||||||
Db        361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIVPAYTMPPNAQHVTVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||||||||: ||| ||:||  |:::    :|:|   :    ||:| :
Db        421 IREDFSRTLAERLVIDIEKVLHELDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

Qy        481 QRDIITGWKKFVAD-RKKTSGIC 502
              | :| | ||||||| :|||:|:|
Db        474 QLEITTAWKKFVADKKKKTNGVC 496

RESULT 3
P93369
ID   P93369      PRELIMINARY;      PRT;   496 AA.
AC   P93369;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Glutamate decarboxylase.
GN   NTGAD1.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Xanthi;
RA   Dharmasiri M.A.N., Lu Y.T., Harrington H.M.;
RT   "Cloning and sequencing of a tobacco cDNA encoding glutamate
RT   decarboxylase.";
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC       TYRDC).
DR   EMBL; U54774; AAB40608.1; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
KW   Decarboxylase; Lyase; Pyridoxal phosphate.
SQ   SEQUENCE   496 AA;  56035 MW;  C48492F4BF0BDA8A CRC64;

  Query Match            85.4%;  Score 2232;  DB 10;  Length 496;
  Best Local Similarity  84.9%;  Pred. No. 1.6e-160;
  Matches  427;  Conservative   35;  Mismatches   33;  Indels    8;  Gaps    3;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   ||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||:||:| |||||||||||||||||||||||||||||||||| : |||||
Db         61 ASFVTTWMEPECNKLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
```

```
b         121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTGANVQVCWEKFARYFEVELKE 180

y         181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
b         181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

y         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              |||||||| |||||||||||||||||||||||||||||||||||||||| |||||||||:||
b         241 IHVDAASGEFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

y         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||:||||||:||  ||||||||
b         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGLEK 360

y         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              : ||||:||: |||||||||||:| | |||||: |||:|||:|||||||||||:||||||
b         361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIIPAYTMPPNAQHVTVLRVV 420

y         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||||||||: ||| ||:||  |:::    :|:|   :    ||:| :
b         421 IREDFSRTLAERLVIDIEKVLHELDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

y         481 QRDIITGWKKFVAD-RKKTSGIC 502
              | :| | ||||||| :|||:|:|
b         474 QLEITTAWKKFVADKKKKTNGVC 496


ESULT 4
9AT17
D   Q9AT17       PRELIMINARY;      PRT;   496 AA.
C   Q9AT17;
T   01-JUN-2001 (TrEMBLrel. 17, Created)
T   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
T   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
E   Glutamate decarboxylase isozyme 3 (EC 4.1.1.15).
S   Nicotiana tabacum (Common tobacco).
C   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
C   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
X   NCBI_TaxID=4097;
N   [1]
P   SEQUENCE FROM N.A.
C   STRAIN=cv. Samsun NN;
A   Yevtushenko D., McLean M.D., Peiris S.E., Van Cauwenberghe O.R.,
A   Shelp B.J.;
T   "Two isoforms of tobacco glutamate decarboxylase are regulated by
T   calcium/calmodulin and differ in organ distribution.";
L   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
C   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
C   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
C       TYRDC).
R   EMBL; AF352732; AAK18620.1; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
W   Decarboxylase; Lyase; Pyridoxal phosphate.
Q   SEQUENCE   496 AA;  55875 MW;  CC61BCD3531827BB CRC64;

 Query Match             85.2%;  Score 2228;  DB 10;  Length 496;
 Best Local Similarity   84.7%;  Pred. No. 3.3e-160;
 Matches  426;  Conservative   35;  Mismatches   34;  Indels    8;  Gaps    3;

y           1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   ||||||:||||||||||||||||||||||||||||||||||||||||||||||
b           1 MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

y          61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||:||:| ||||||||| |||||||||||||||||||||||| : |||||
b          61 ASFVTTWMEPECNKLMMDSINKNYVDMGEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

y         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
b         121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTGANVQVCWEKFARYFEVELKE 180
```

```
Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
Db        181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||:||
Db        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||:||||||:||  ||||||||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGLEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              : ||||:||: |||||||||||:| | |||||: |||:|||:|||||||||||:||||||
Db        361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIIPAYTMPPNAQHVTVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||||||||: ||| ||:||  |:::    :|:|   :    ||:| :
Db        421 IREDFSRTLAERLVIDIEKVLHELDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

Qy        481 QRDIITGWKKFVAD-RKKTSGIC 502
              | :|   ||||||| :|||:|:|
Db        474 QLEITAAWKKFVADKKKKTNGVC 496


RESULT 5
O81101
ID   O81101       PRELIMINARY;      PRT;   496 AA.
AC   O81101;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Glutamate decarboxylase isozyme 2.
GN   NTGAD2.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98302498; PubMed=9638642;
RA   Yun S.J., Oh S.H.;
RT   "Cloning and characterization of a tobacco cDNA encoding
RT   calcium/calmodulin-dependent glutamate decarboxylase.";
RL   Mol. Cells 8:125-129(1998).
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC       TYRDC).
DR   EMBL; AF020424; AAC39483.1; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
KW   Decarboxylase; Lyase; Pyridoxal phosphate.
SQ   SEQUENCE   496 AA;  55931 MW;  5B9D2C8E12560D27 CRC64;

  Query Match             84.1%;  Score 2199;  DB 10;  Length 496;
  Best Local Similarity   84.1%;  Pred. No. 5.2e-158;
  Matches  423;  Conservative   34;  Mismatches   38;  Indels    8;  Gaps    3;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVLSKTASESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||: |:| |||||||||||||||||||||||||||||||||| : |||||
Db         61 ASFVTTWMEPECNTLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPFDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
```

```
)         181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

y         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||| ||||||||||||||||||| |||||||||:||
)         241 IHVDAASGGFIAPFLYPELEWDFRLPLEKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

y         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||:||||||:||  |||||:||
)         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGIEK 360

y         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              : ||||:||: |||||||||||:| | |||||: |||:|||| ||||||||||:||||||
)         361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIVLAYTMPPNAQHVTVLRVV 420

y         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||||||||   :| ||:||  |:::    :|:|   :    ||:| : 
)         421 IREDFSRTLAERLVIDIEKVFHGVDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

y         481 QRDIITGWKKFVAD-RKKTSGIC 502
              | :| | | ||||| :|||:|:|
)         474 QLEITTAWLKFVADKKKKTNGVC 496


ESULT 6
8LLP2
D   Q8LLP2      PRELIMINARY;      PRT;   513 AA.
C   Q8LLP2;
T   01-OCT-2002 (TrEMBLrel. 22, Created)
T   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
T   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E   Putative glutamate carboxylase.
N   OSJNBA0031O09.06.
S   Oryza sativa (japonica cultivar-group).
C   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
C   Ehrhartoideae; Oryzeae; Oryza.
X   NCBI_TaxID=39947;
N   [1]
P   SEQUENCE FROM N.A.
A   Eastman A.P., Smith S.C., Bertin N., Liang C., Gingle A.R.,
A   Pratt L.H., Cordonnier-Pratt M.-M.;
T   "Sequencing of clone OSJNBa0031O09.";
L   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
N   [2]
P   SEQUENCE FROM N.A.
A   Eastman A.P., Smith S.C., Bertin N., Liang C., Najar F.Z., Pratt L.H.,
A   Cordonnier-Pratt M.-M.;
T   "Untitled.";
L   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
C   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
C   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
C       TYRDC).
R   EMBL; AF377946; AAM47304.1; -.
R   Gramene; Q8LLP2; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
W   Decarboxylase; Lyase; Pyridoxal phosphate.
Q   SEQUENCE   513 AA;  57378 MW;  131DFD62F6E0CE87 CRC64;

 Query Match             83.8%;  Score 2191.5;  DB 10;  Length 513;
 Best Local Similarity   79.7%;  Pred. No. 2e-157;
 Matches  417;  Conservative   38;  Mismatches   37;  Indels   31;  Gaps    2;

)y          1 MVLSHAVSESDVSVHSTFASRYVRTSLPR---------------------FKMPENSIPK 39
              ||||| || || |||||||||||||||||                     |:||| ||||
)b          1 MVLSHGVSGSDESVHSTFASRYVRTSLPRHARSPLSRAPLAPIDSVIDWEFRMPEQSIPK 60

)y         40 EAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNR 99
              ||||||||||||||||||||||||||||||||||||| :|:|||||||||||||||||||
)b         61 EAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIQASVNKNYVDMDEYPVTTELQNR 120

)y        100 CVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVT 159
              ||||||||||||| ::|||||||||||||||||||||||||:|||| || ||| |||||||
```

```
Db        121 CVNMIAHLFNAPLGDSETAVGVGTVGSSEAIMLAGLAFKRRWQNKMKAAGKPCDKPNIVT 180

Qy        160 GANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFE 219
              |||||||||||||||||||||||||:||||||| :||||||||||||| |||||||||||
Db        181 GANVQVCWEKFARYFEVELKEVKLSDGYYVMDPAKAVDMVDENTICVAAILGSTLNGEFE 240

Qy        220 DVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYG 279
              ||||||||| :|| ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 DVKLLNDLLTKKNAETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYG 300

Qy        280 LVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEG 339
              |||||||| |||:|||||||||||||||||||||||||||||||||||||||||||| ||
Db        301 LVYAGIGWCIWRSKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEG 360

Qy        340 YRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYG 399
              |:|:||||:|| :||::||||| |||||||| ||||||||||||: | |||||| |||:|
Db        361 YKNIMENCQENAMVLKQGLEKTGRFNIVSKDNGVPLVAFSLKDSARHNEFEISDFLRRFG 420

Qy        400 WIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQ 459
              ||||||||||:|||:||||||||||||||||||||:|:|||: ||| ||:||:       
Db        421 WIVPAYTMPPDAQHVTVLRVVIREDFSRTLAERLVLDVEKVLHELDALPARVV------- 473

Qy        460 EKSESNSDNLMVTVKKSDIDKQRDIITGWKKFVADRKKTSGIC 502
                  :  |    :  : :::|||::|:  ||: |  :|||:|:|
Db        474 ---ANGGDAAAASASEREMEKQREVISLWKRAVLAKKKTNGVC 513


RESULT 7
Q94KK8
ID   Q94KK8      PRELIMINARY;      PRT;   491 AA.
AC   Q94KK8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Glutamate decarboxylase isozyme 4 (EC 4.1.1.15).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yevtushenko D., McLean M.D., Peiris S.E., Van Cauwenberghe O.R.,
RA   Shelp B.J.;
RT   "Two isoforms of tobacco glutamate decarboxylase are regulated by
RT   calcium/calmodulin and differ in organ distribution.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC       TYRDC).
DR   EMBL; AF353615; AAK38667.1; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
KW   Decarboxylase; Lyase; Pyridoxal phosphate.
SQ   SEQUENCE   491 AA;  55950 MW;  51CB78B3EF3F4E35 CRC64;

  Query Match             83.6%;  Score 2186.5;  DB 10;  Length 491;
  Best Local Similarity   84.9%;  Pred. No. 4.5e-157;
  Matches  421;  Conservative   27;  Mismatches   39;  Indels    9;  Gaps    1;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   ||||||||||||||||||||||| :| |||||||||:|||||||||||||||||
Db          1 MVLSKTSSESDVSVHSTFASRYVRTSLPRFEMAENSIPKEAAFQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              |||||||||||||||:| ||||||||||||||||||||||||||| |||||||| |||||
Db         61 ASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVNMIARLFNAPLEEKETAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||||||||||||||| ||||||||||| :|||||||||||||||||| ||||||||
Db        121 VGTVGSSEAIMLAGLAFKRNWQNKRKAEGKPYNKPNIVTGANVQVCWEKFANYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
```

```
              ||| |||||||| |||:|||||||||| ||||||||||||||||||||:||||:|||:||
b         181 VKLREGYYVMDPVQAVEMVDENTICVAAILGSTLNGEFEDVKLLNDLLIEKNKQTGWNTP 240

y         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |:||||||
b         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRTKQDLPEEL 300

y         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              ||||||||||||||||||||||||||||||||||||:||||||||||||| |||||||||
b         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGYEGYRNVMENCRENAIVLREGLEK 360

y         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | ||||||||||||||||||||:| | |||:|: |||:|||||||||| :|||:||||||
b         361 TGRFNIVSKDEGVPLVAFSLKDNSRHNEFEVSETLRRFGWIVPAYTMPADAQHVTVLRVV 420

y         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||:|| ||: ||| ||:|:  |:             |:   || :::
b         421 IREDFSRTLAERLVLDIVKVLHELDTLPARLSAKL---------EEVKLVKNGKKFELEV 471

y         481 QRDIITGWKKFVADRK 496
              ||::   |||||  ||
b         472 QREVTNYWKKFVLARK 487


ESULT 8
9ZPS3
D    Q9ZPS3       PRELIMINARY;      PRT;   493 AA.
C    Q9ZPS3;
T    01-MAY-1999 (TrEMBLrel. 10, Created)
T    01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
T    01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
E    Putative glutamate decarboxylase (AT2G02010/F14H20.8).
N    F14H20.8.
S    Arabidopsis thaliana (Mouse-ear cress).
C    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
C    eurosids II; Brassicales; Brassicaceae; Arabidopsis.
X    NCBI_TaxID=3702;
N    [1]
P    SEQUENCE FROM N.A.
C    STRAIN=cv. Columbia;
A    Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
A    Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.,
A    Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
A    Fraser C.M., Venter J.C.;
T    "Arabidopsis thaliana chromosome II BAC F14H20 genomic sequence.";
L    Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
N    [2]
P    SEQUENCE FROM N.A.
A    Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
A    Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
A    Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,
A    Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
A    Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,
A    Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W.,
A    Theologis A., Ecker J.R.;
T    "Arabidopsis cDNA clones.";
L    Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
N    [3]
P    SEQUENCE FROM N.A.
A    Cheuk R., Chen H., Kim C.J., Shinn P., Banh J., Bowser L.,
A    Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
A    Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
A    Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
A    Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
A    Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
A    Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
T    "Arabidopsis ORF clones.";
L    Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
C    -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
C    -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
C        TYRDC).
R    EMBL; AC006532; AAD20099.1; -.
```

```
DR   EMBL; AF361836; AAK32848.1; -.
DR   EMBL; AY124860; AAM70569.1; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
KW   Decarboxylase; Lyase; Pyridoxal phosphate.
SQ   SEQUENCE   493 AA;  56005 MW;  36B3FC93F2978168 CRC64;

  Query Match             81.7%;  Score 2135.5;  DB 10;  Length 493;
  Best Local Similarity   82.4%;  Pred. No. 3.3e-153;
  Matches  411;  Conservative   32;  Mismatches   45;  Indels   11;  Gaps    2;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||  |||||||:|||||||||| |||||:|||||||||||||||||||||||||||||
Db          1 MVLSKTVSESDVSIHSTFASRYVRNSLPRFEMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              |||||||||||||||:| ||||||||||||||||||||||||||| |||||| : | |||
Db         61 ASFVTTWMEPECDKLMMESINKNYVDMDEYPVTTELQNRCVNMIARLFNAPLGDGEAAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||:|||||||:| | ||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRQWQNKRKAQGLPYDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              | | | |||||| :||:|||||||||| |||||| ||||||||||||||||||:||||||
Db        181 VNLREDYYVMDPVKAVEMVDENTICVAAILGSTLTGEFEDVKLLNDLLVEKNKQTGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||:|| | |||:||
Db        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | || ||||: ||||||||||||| | |||::  |||:|||||||||| :|||:||||||
Db        361 TGRFKIVSKENGVPLVAFSLKDSSRHNEFEVAHTLRRFGWIVPAYTMPADAQHVTVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| | |||: ||| ||:||  |:: |:             |||:  :
Db        421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMANGKVNG----------VKKTPEET 470

Qy        481 QRDIITGWKKFVADRKKTS 499
              ||::   ||| : : |||:
Db        471 QREVTAYWKKLL-ETKKTN 488


RESULT 9
Q944L6
ID   Q944L6       PRELIMINARY;      PRT;   494 AA.
AC   Q944L6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   At1g65960/F12P19_12.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA   Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA   Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA   Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA   Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA   Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA   Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA   Ecker J.R.;
RT   "Arabidopsis cDNA clones.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
```

```
C   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
C   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
C       TYRDC).
R   EMBL; AF428294; AAL16126.1; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
W   Decarboxylase; Lyase; Pyridoxal phosphate.
Q   SEQUENCE   494 AA;  56168 MW;  9308E4A2456423C4 CRC64;

 Query Match              81.3%;  Score 2126;  DB 10;  Length 494;
 Best Local Similarity    80.0%;  Pred. No. 1.7e-152;
 Matches  403;  Conservative   46;  Mismatches   43;  Indels   12;  Gaps    3;

y           1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              |||:   : :| || : | ||||||:||:::: ||||||:|||||| |||||||||||||
b           1 MVLTKTAT-NDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQIIKDELMLDGNPRLNL 59

y          61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||| ||||||||||||||||||||||||:|| ||||||||:|||||
b          60 ASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNIIARLFNAPLEESETAVG 119

y         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
b         120 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGANVQVCWEKFARYFEVELKE 179

y         181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              | |||||||||| :| :|||||||||| |||||||||||||| ||||||:||:||||:||
b         180 VNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVKKNEETGWNTP 239

y         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||:|||||||||||||||||||||||||||||||||:||  |||||||
b         240 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRAAEDLPEEL 299

y         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              ||||||||||||||||||||||||:||||||||||| |||:|||||| |||:||:||:||
b         300 IFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRLGFEGYKNVMENCIENMVVLKEGIEK 359

y         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              ||||||||||:|||:||||||| | | |||||:||||:|||||||||| : |||||||||
b         360 TERFNIVSKDQGVPVVAFSLKDHSFHNEFEISEMLRRFGWIVPAYTMPADVQHITVLRVV 419

y         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| || ||: ||| |||::  |: :         : :   ||:  ::|
b         420 IREDFSRTLAERLVADISKVLHELDTLPSKISKKMGI---------EGIAENVKEKKMEK 470

y         481 Q--RDIITGWKKFVADRKKTSGIC 502
              :   ::| ||:||| :||| :|:|
b         471 EILMEVIVGWRKFVKERKKMNGVC 494

ESULT 10
9ZPS4
D   Q9ZPS4      PRELIMINARY;      PRT;   500 AA.
C   Q9ZPS4;
T   01-MAY-1999 (TrEMBLrel. 10, Created)
T   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
T   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E   Putative glutamate decarboxylase.
N   AT2G02000 OR AT2G02000/F14H20.7.
S   Arabidopsis thaliana (Mouse-ear cress).
C   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=cv. Columbia;
RX  MEDLINE=20083487; PubMed=10617197;
RA  Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA  Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA  Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA  Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA  Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
```

```
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA   Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA   Hayashizaki Y., Shinozaki K.;
RT   "Arabidopsis thaliana full-length cDNA.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC       TYRDC).
DR   EMBL; AC006532; AAD20093.1; -.
DR   EMBL; AK118125; BAC42751.1; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
KW   Decarboxylase; Lyase; Pyridoxal phosphate.
SQ   SEQUENCE   500 AA;  56506 MW;  1A47C0D14C6CD3B8 CRC64;

  Query Match              79.5%;  Score 2080;  DB 10;  Length 500;
  Best Local Similarity    78.8%;  Pred. No. 5.4e-149;
  Matches  395;  Conservative   44;  Mismatches   52;  Indels   10;  Gaps    1;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   |:|| |:|||||||||| |: ||::|:||||||||||||||||  ||||||||
Db          1 MVLSKTASKSDDSIHSTFASRYVRNSISRFEIPKNSIPKEAAYQIINDELKFDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              |||||||||||||||:| ||||| |:||:|||||:|||||||||| |||||| : | |:|
Db         61 ASFVTTWMEPECDKLMMESINKNNVEMDQYPVTTDLQNRCVNMIARLFNAPLGDGEAAIG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||:|||||||||:||||||| | | |:||||||||:||| |||||||||||||
Db        121 VGTVGSSEAVMLAGLAFKRQWQNKRKALGLPYDRPNIVTGANIQVCLEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||| |||||||| :||:|||||||||  |||||| ||||||||||||||||||:||||||
Db        181 VKLREGYYVMDPDKAVEMVDENTICVVAILGSTLTGEFEDVKLLNDLLVEKNKKTGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              |||||||||||||||||:|||||||||||||||||||||||||||||||:|| | |||:||
Db        241 IHVDAASGGFIAPFLYPDLEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | |||||||: ||||||||||||| | |||:::||||:|||||||||| :|||:||||||
Db        361 TGRFNIVSKENGVPLVAFSLKDSSRHNEFEVAEMLRRFGWIVPAYTMPADAQHVTVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| | |||: ||| ||:||  |:: |:             |||:  : 
Db        421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMASGKVNG----------VKKTPEET 470

Qy        481 QRDIITGWKKFVADRKKTSGI 501
              ||::   |||||  :   :|:
Db        471 QREVTAYWKKFVDTKTDKNGV 491

RESULT 11
Q9AQU4
ID   Q9AQU4      PRELIMINARY;      PRT;   501 AA.
AC   Q9AQU4;
```

```
T   01-JUN-2001 (TrEMBLrel. 17, Created)
T   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
T   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
E   Glutamate decarboxylase (EC 4.1.1.15).
N   GAD.
S   Oryza sativa (Rice).
C   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
C   Ehrhartoideae; Oryzeae; Oryza.
X   NCBI_TaxID=4530;
N   [1]
P   SEQUENCE FROM N.A.
C   STRAIN=cv. Nipponbare; TISSUE=Shoot;
A   Akama K., Akihiro T., Kitagawa M., Takaiwa F.;
T   "Molecular characterization of two genes encoding glutamate
T   decarboxylase from rice (Oryza sativa).";
L   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
C   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
C   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
C       TYRDC).
R   EMBL; AB056062; BAB32870.1; -.
R   EMBL; AB056060; BAB32868.1; -.
R   Gramene; Q9AQU4; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
W   Decarboxylase; Lyase; Pyridoxal phosphate.
Q   SEQUENCE   501 AA;  56657 MW;  EB4962C147FFA7A8 CRC64;

 Query Match             78.6%;  Score 2056.5;  DB 10;  Length 501;
 Best Local Similarity   78.1%;  Pred. No. 3.2e-147;
 Matches  395;  Conservative   44;  Mismatches   58;  Indels    9;  Gaps    4;

y           1 MVLSHAVSESD----VSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNP 56
              ||:| | ::||    |   : |||||||  ||||:||| |||:|||||||||||||||||
b           1 MVVSVAATDSDTAQPVQYSTFFASRYVRDPLPRFRMPEQSIPREAAYQIINDELMLDGNP 60

y          57 RLNLASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAE 116
              ||||||||||||||||||||| |:|||||||||||||||||||||||||||||||::| |
b          61 RLNLASFVTTWMEPECDKLIMDSVNKNYVDMDEYPVTTELQNRCVNMIAHLFNAPIKEDE 120

y         117 TAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEV 176
              ||:||||||||||||||||||||||||||| :||| ||||||||||||||||||||||||
b         121 TAIGVGTVGSSEAIMLAGLAFKRKWQNKRKEQGKPCDKPNIVTGANVQVCWEKFARYFEV 180

y         177 ELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETG 236
              |||||||||||||||| :||:|||||||||| |||||| ||||||||||:|| |||||||
b         181 ELKEVKLSEGYYVMDPVKAVEMVDENTICVAAILGSTLTGEFEDVKLLNNLLTEKNKETG 240

y         237 WDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDL 296
              || ||||||||||||||||||||||||||||||||||||||||||| |:||||||:||||
b         241 WDVPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYPGVGWVIWRSKEDL 300

y         297 PEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLRE 356
              |||||||||||| |||||||||||||||:||||||||||| |||:|:|:|| ||  :|||
b         301 PEELIFHINYLGTDQPTFTLNFSKGSSQIIAQYYQLIRLGFEGYKNIMQNCMENTAILRE 360

y         357 GLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITV 416
              |:| | || |:||: ||||||||||||  :| |:||: |||:|||||||||| ||:|: |
b         361 GIEATGRFEILSKEAGVPLVAFSLKDSGRYTVFDISEHLRRFGWIVPAYTMPANAEHVAV 420

y         417 LRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKS 476
              |||||||||||:|||||| || |:: |||   ::|: |||    | :|  |   |  |||
b         421 LRVVIREDFSRSLAERLVSDIVKILHELDAHSAQVL-KISSAIAKQQSGDDG--VVTKKS 477

y         477 DIDKQRDIITGWKKFVADRKKTSGIC 502
               :: :|:|   |:  |  :|| :|||
b         478 VLETEREIFAYWRDQV--KKKQTGIC 501


ESULT 12
9LSH2
D   Q9LSH2      PRELIMINARY;      PRT;   494 AA.
C   Q9LSH2;
```

```
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Glutamate decarboxylase.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RA   Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Columbia;
RX   MEDLINE=20277480; PubMed=10819329;
RA   Nakamura Y.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT   features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT   clones.";
RL   DNA Res. 7:131-135(2000).
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC       TYRDC).
DR   EMBL; AB026646; BAB02870.1; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
KW   Decarboxylase; Lyase; Pyridoxal phosphate.
SQ   SEQUENCE   494 AA;  55770 MW;  7985F175E54DF262 CRC64;

  Query Match             75.3%;  Score 1969;  DB 10;  Length 494;
  Best Local Similarity   75.2%;  Pred. No. 1.3e-140;
  Matches  373;  Conservative   56;  Mismatches   61;  Indels    6;  Gaps    3;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              |||:   |:||  :||||||||||  :||||||:: :||:||||:|||||||||||||||
Db          1 MVLA-TNSDSDEHLHSTFASRYVRAVVPRFKMPDHCMPKDAAYQVINDELMLDGNPRLNL 59

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||| |:|||||||||||||||||||||||||:||:||: | | |:|
Db         60 ASFVTTWMEPECDKLIMDSVNKNYVDMDEYPVTTELQNRCVNMIANLFHAPVGEDEAAIG 119

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
               |||||||||||||||||||||::|||:| |:||||||||||||||||||||||||||||
Db        120 CGTVGSSEAIMLAGLAFKRKWQHRRKAQGLPIDKPNIVTGANVQVCWEKFARYFEVELKE 179

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||| |||||| :||:|||||||||| |||||| ||||||| ||||| ||| ||||:||
Db        180 VKLSEDYYVMDPAKAVEMVDENTICVAAILGSTLTGEFEDVKQLNDLLAEKNAETGWETP 239

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              |||||||||||||||||:|||||||| |||||||||||||||||:|||:||:|:||||||
Db        240 IHVDAASGGFIAPFLYPDLEWDFRLPWVKSINVSGHKYGLVYAGVGWVVWRTKDDLPEEL 299

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              :|||||||||||||||||||||||:|||||| |||| |||:|:|||| :|   ||||:|
Db        300 VFHINYLGADQPTFTLNFSKGSSQIIAQYYQFIRLGFEGYKNIMENCMDNARRLREGIEM 359

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | ::||||||| ||||||||||||| || |||:: ||::|||:|||||| :|||| |||||
Db        360 TGKFNIVSKDIGVPLVAFSLKDSSKHTVFEIAESLRKFGWIIPAYTMPADAQHIAVLRVV 419

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||| ||:||:  | :|::|:: ||||: |  :     :  : |:  | || : :
Db        420 IREDFSRGLADRLITHIIQVLKEIEGLPSRIAHLAA----AAAVSGDDEEVKVKTAKMSL 475

Qy        481 QRDIITGWKKFVADRK 496
              : ||   ||: |  ::
Db        476 E-DITKYWKRLVEHKR 490
```

```
ESULT 13
8LFR4
D   Q8LFR4      PRELIMINARY;      PRT;   494 AA.
C   Q8LFR4;
T   01-OCT-2002 (TrEMBLrel. 22, Created)
T   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
T   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
E   Glutamate decarboxylase, putative.
S   Arabidopsis thaliana (Mouse-ear cress).
C   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
C   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
C   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
X   NCBI_TaxID=3702;
N   [1]
P   SEQUENCE FROM N.A.
A   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
A   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
T   "Full-length messenger RNA sequences greatly improve genome
T   annotation.";
L   Genome Biol. 0:0-0(2002).
N   [2]
P   SEQUENCE FROM N.A.
A   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
A   Feldmann K.;
T   "Full-Length cDNA from Arabidopsis thaliana.";
L   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
C   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
C   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
C       TYRDC).
R   EMBL; AY084689; AAM61251.1; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
W   Decarboxylase; Lyase; Pyridoxal phosphate.
Q   SEQUENCE   494 AA;  55804 MW;  7B875CB82EC23E20 CRC64;

 Query Match             75.1%;  Score 1965;  DB 10;  Length 494;
 Best Local Similarity   75.0%;  Pred. No. 2.7e-140;
 Matches  372;  Conservative   56;  Mismatches   62;  Indels    6;  Gaps    3;

y            1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
               |||:   |:||  :||||||||||  :||||||::  :||:||||:|||||||||||||||
b            1 MVLA-TNSDSDEHLHSTFASRYVRAVVPRFKMPDHCMPKDAAYQVINDELMLDGNPRLNL 59

y           61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
               ||||||||||||||||| |:|||||||||||||||||||||||||: |:||: | | |:|
b           60 ASFVTTWMEPECDKLIMDSVNKNYVDMDEYPVTTELQNRCVNMIANFFHAPVGEDEAAIG 119

y          121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
                |||||||||||||||||||||::|||:| |:||||||||||||||||||||||||||||
b          120 CGTVGSSEAIMLAGLAFKRKWQHRRKAQGLPIDKPNIVTGANVQVCWEKFARYFEVELKE 179

y          181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
               ||||| |||||| :||:|||||||||| |||||| ||||||| ||||| ||| ||||:||
b          180 VKLSEDYYVMDPAKAVEMVDENTICVAAILGSTLTGEFEDVKQLNDLLAEKNAETGWETP 239

y          241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
               |||||||||||||||||:|||||||| |||||||||||||||||:|||:|| |:||||||
b          240 IHVDAASGGFIAPFLYPDLEWDFRLPWVKSINVSGHKYGLVYAGVGWVVWRTKDDLPEEL 299

y          301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
               :|||||||||||||||||||||||:|||||| |||| |||:|:|||| :|   ||||:| 
Db         300 VFHINYLGADQPTFTLNFSKGSSQIIAQYYQFIRLGFEGYKNIMENCMDNARRLREGIEM 359

Qy         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
               | :||||||| ||||||||||||| || |||:: ||::|||:|||||| :|||| |||||
Db         360 TGKFNIVSKDIGVPLVAFSLKDSSKHTVFEIAESLRKFGWIIPAYTMPADAQHIAVLRVV 419

Qy         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
               ||||||| ||:||:  | :|::|:: ||||: |  :     :  : |:  | || : :
Db         420 IREDFSRGLADRLITHIIQVLKEIEGLPSRIAHLAA----AAAVSGDDEEVKVKTAKMSL 475

Qy         481 QRDIITGWKKFVADRK 496
               : ||   ||: |  ::
```

```
Db         476 E-DITKYWKRLVEHKR 490

RESULT 14
Q8RXH0
ID   Q8RXH0      PRELIMINARY;      PRT;   419 AA.
AC   Q8RXH0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Glutamate decarboxylase, putative.
GN   AT1G65960.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA   Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA   Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA   Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA   Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA   Ecker J., Theologis A., Davis R.W.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC       TYRDC).
DR   EMBL; AY081259; AAL91148.1; -.
DR   InterPro; IPR002129; Pyridoxal_deC.
DR   Pfam; PF00282; pyridoxal_deC; 1.
KW   Decarboxylase; Lyase; Pyridoxal phosphate.
SQ   SEQUENCE   419 AA;  47639 MW;  7565ADB31A76957F CRC64;

  Query Match             70.6%;  Score 1845.5;  DB 10;  Length 419;
  Best Local Similarity   81.3%;  Pred. No. 2.3e-131;
  Matches  348;  Conservative   36;  Mismatches   33;  Indels   11;  Gaps    2;

Qy          77 MSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLA 136
               | ||||||||||||||||||||||||:|| ||||||||:|||||||||||||||||||||
Db           1 MDSINKNYVDMDEYPVTTELQNRCVNIIARLFNAPLEESETAVGVGTVGSSEAIMLAGLA 60

Qy         137 FKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAV 196
               ||||||||||||||| ||||||||||||||||||||||||||||| |||||||||| :| 
Db          61 FKRKWQNKRKAEGKPYDKPNIVTGANVQVCWEKFARYFEVELKEVNLSEGYYVMDPDKAA 120

Qy         197 DMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLY 256
               :|||||||||| |||||||||||||| ||||||:||:||||:||||||||||||||||:|
Db         121 EMVDENTICVAAILGSTLNGEFEDVKRLNDLLVKKNEETGWNTPIHVDAASGGFIAPFIY 180

Qy         257 PELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTL 316
               ||||||||||||||||||||||||||||||||:||  |||||||||||||||||||||||
Db         181 PELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRAAEDLPEELIFHINYLGADQPTFTL 240

Qy         317 NFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLV 376
               ||||||||:||||||||||| |||:|||||| |||:||:||:||||||||||||:|||:|
Db         241 NFSKGSSQIIAQYYQLIRLGFEGYKNVMENCIENMVVLKEGIEKTERFNIVSKDQGVPVV 300

Qy         377 AFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVID 436
               |||||| | | |||||:||||:|||||||||| :|||||||||||||||||||||||| |
Db         301 AFSLKDHSFHNEFEISEMLRRFGWIVPAYTMPADAQHITVLRVVIREDFSRTLAERLVAD 360

Qy         437 IEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDKQ--RDIITGWKKFVAD 494
               | ||: ||| |||::  |: :         : :   ||:  ::|:   ::| ||:||| :
Db         361 ISKVLHELDTLPSKISKKMGI---------EGIAKNVKEKKMEKEILMEVIVGWRKFVKE 411

Qy         495 RKKTSGIC 502
               ||| :|:|
Db         412 RKKMNGVC 419

RESULT 15
```

```
AR41
)   Q9AR41      PRELIMINARY;      PRT;   500 AA.
:   Q9AR41;
`   01-JUN-2001 (TrEMBLrel. 17, Created)
`   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
`   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
:   Glutamate decarboxylase (EC 4.1.1.15).
1   GAD.
,   Oryza sativa (Rice).
:   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
:   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
:   Ehrhartoideae; Oryzeae; Oryza.
(   NCBI_TaxID=4530;
1   [1]
>   SEQUENCE FROM N.A.
:   STRAIN=cv. Nipponbare; TISSUE=Shoot;
A   Akama K., Akihiro T., Kitagawa M., Takaiwa F.;
Γ   "Molecular characterization of two genes encoding glutamate
Γ   decarboxylase from rice (Oryza sativa).";
J   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
:   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
:   -!- SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
:       TYRDC).
R   EMBL; AB056063; BAB32871.1; -.
R   EMBL; AB056061; BAB32869.1; -.
R   Gramene; Q9AR41; -.
R   InterPro; IPR002129; Pyridoxal_deC.
R   Pfam; PF00282; pyridoxal_deC; 1.
N   Decarboxylase; Lyase; Pyridoxal phosphate.
)   SEQUENCE   500 AA;  55615 MW;  0E19E61CCAD5F508 CRC64;

 Query Match             70.5%;  Score 1844;  DB 10;  Length 500;
 Best Local Similarity   69.6%;  Pred. No. 3.9e-131;
 Matches  355;  Conservative   56;  Mismatches   81;  Indels   18;  Gaps    5;

y          1 MVLSH--AVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRL 58
             |||:|  || |   :  : ||||||:  :||::: | || |:|||||::|||:|| :|||
b          1 MVLTHVEAVEEGSEAAAAVFASRYVQDPVPRYELGERSISKDAAYQIVHDELLLDSSPRL 60

y         59 NLASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETA 118
             |||||||||||||||:||: :||||| |||||||||||||||||:|| |||||: : | |
b         61 NLASFVTTWMEPECDRLILEAINKNYADMDEYPVTTELQNRCVNIIARLFNAPVGDGEKA 120

y        119 VGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVEL 178
             |||||||||||||||||||||:|||:||| ||| ||||||||||||||||||||||||||
b        121 VGVGTVGSSEAIMLAGLAFKRRWQNRRKAAGKPHDKPNIVTGANVQVCWEKFARYFEVEL 180

y        179 KEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWD 238
             |||||:|| ||||| :||||||||||||| |||||| ||||||: |||||  ||| ||||
b        181 KEVKLTEGCYVMDPVKAVDMVDENTICVAAILGSTLTGEFEDVRRLNDLLAAKNKRTGWD 240

y        239 TPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPE 298
             ||||||||||||||||:|||||||||||||||||||||||||||||:|||||||||||||
b        241 TPIHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGVGWVIWRNKEDLPE 300

y        299 ELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGL 358
             ||||||||||||||||||||||||||:|||||| :||| |||::||:|| |:   |||||
b        301 ELIFHINYLGADQPTFTLNFSKGSSQIIAQYYQFLRLGFEGYKSVMKNCMESARTLREGL 360

)y       359 EKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLR 418
             ||| || |:||:||||||||: || :    | :|  ||||||||||||||   :|:||:|
)b       361 EKTGRFTIISKEEGVPLVAFTFKDGAGAQAFRLSSGLRRYGWIVPAYTMPAALEHMTVVR 420

Qy       419 VVIREDFSRTLAERLVIDIEKVMRELD-----ELPSRVIHKISLGQEKSESNSDNLMVTV 473
             ||:|||| | |||| :  :   : |:|      :| ||   | ||  ::     :: |
)b       421 VVVREDFGRPLAERFLSHVRMALDEMDLAARAPVP-RVQLTIELGPARTAGEEASIRVVK 479

Qy       474 KKS-DIDKQRDIITGWKKFVADRKKTSGIC 502
              ::  : |   :: |         || |:|
)b       480 SEAVPVRKSVPLVAG---------KTKGVC 500

Search completed: October 22, 2003, 15:07:41
```

Job time : 73 secs

```
                      GenCore version 5.1.6
             Copyright (c) 1993 - 2003  Compugen Ltd.

M protein - nucleic search, using frame_plus_p2n model

un on:          October 22, 2003, 16:47:21 ; Search time 359 Seconds
                                           (without alignments)
                                           3750.039 Million cell updates/sec

itle:           US-10-006-852-2
erfect score:   2615
equence:        1 MVLSHAVSESDVSVHSTFAS.........DIITGWKKFVADRKKTSGIC 502

coring table:   BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

earched:        1792395 seqs, 1340900451 residues

otal number of hits satisfying chosen parameters:      3584790

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

ommand line parameters:
MODEL=frame+_p2n.model -DEV=xlp
Q=/cgn2_1/USPTO_spool_p/US10006852/runat_21102003_164105_2567/app_query.fasta_1.647
DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
MAXLEN=2000000000 -USER=US10006852_@CGN_1_1_107_@runat_21102003_164105_2567
NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :       Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                16:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                17:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                %
esult           Query
 No.    Score   Match Length DB  ID                           Description
-----------------------------------------------------------------------------
   1     2615   100.0    1509  14  US-10-006-852-1            Sequence 1, Appli
```

```
      2    2607   99.7   1509  10  US-09-938-842A-1810        Sequence 1810, Ap
      3    2250   86.0   1785  14  US-10-006-852-15           Sequence 15, Appl
      4    2240   85.7   1705  14  US-10-006-852-11           Sequence 11, Appl
      5    2233   85.4   1745  14  US-10-005-602-1            Sequence 1, Appli
      6    2200   84.1   1479  10  US-09-887-576-820          Sequence 820, App
      7    2199   84.1   1771  14  US-10-006-852-13           Sequence 13, Appl
      8  2135.5   81.7   1482  10  US-09-938-842A-937         Sequence 937, App
      9    2130   81.5   1665  14  US-10-006-852-3            Sequence 3, Appli
     10    2073   79.3   1497  12  US-10-167-547C-1           Sequence 1, Appli
     11    2014   77.0   1783  14  US-10-006-852-17           Sequence 17, Appl
     12    2007   76.7   1455  10  US-09-887-576-782          Sequence 782, App
     13  1936.5   74.1   2121  14  US-10-006-852-7            Sequence 7, Appli
     14  1889.5   72.3   1287  10  US-09-887-576-806          Sequence 806, App
     15    1864   71.3   1530  12  US-10-167-547C-3           Sequence 3, Appli
     16    1864   71.3   1590  12  US-10-167-547C-5           Sequence 5, Appli
     17    1854   70.9   1467  10  US-09-887-576-777          Sequence 777, App
     18    1836   70.2   1180  11  US-09-727-758-1            Sequence 1, Appli
     19  1834.5   70.2   2493  14  US-10-006-852-5            Sequence 5, Appli
     20    1782   68.1   1946  14  US-10-006-852-9            Sequence 9, Appli
     21    1630   62.3   5160  10  US-09-887-576-443          Sequence 443, App
     22  1314.5   50.3   1302  10  US-09-887-576-783          Sequence 783, App
     23    1184   45.3    894  12  US-10-167-547C-30          Sequence 30, Appl
     24    1082   41.4   1410  14  US-10-156-761-4639         Sequence 4639, Ap
c    25    1082   41.4 9025608  14  US-10-156-761-1            Sequence 1, Appli
     26    1051   40.2   1362  14  US-10-156-761-3588         Sequence 3588, Ap
     27    1051   40.2 9025608  14  US-10-156-761-1            Sequence 1, Appli
     28     934   35.7    742  10  US-09-966-881-20           Sequence 20, Appl
     29     851   32.5    807  10  US-09-966-881-21           Sequence 21, Appl
     30   620.5   23.7    440  9  US-09-770-444-922          Sequence 922, App
     31     545   20.8    387  10  US-09-878-574-1480         Sequence 1480, Ap
c    32     452   17.3    493  11  US-09-770-961-610          Sequence 610, App
     33     438   16.7    280  10  US-09-878-574-7302         Sequence 7302, Ap
     34   436.5   16.7    781  11  US-09-989-442-73           Sequence 73, Appl
     35     347   13.3    277  10  US-09-878-574-13198        Sequence 13198, A
     36     325   12.4   1770  12  US-10-197-073-7            Sequence 7, Appli
     37     325   12.4   1770  12  US-10-053-510-1            Sequence 1, Appli
     38     325   12.4   1770  14  US-10-286-175-7            Sequence 7, Appli
c    39   315.5   12.1    477  9  US-09-770-444-28           Sequence 28, Appl
     40   299.5   11.5   1638  12  US-10-053-510-15           Sequence 15, Appl
     41   265.5   10.2   1629  12  US-10-197-073-5            Sequence 5, Appli
     42   265.5   10.2   1629  12  US-10-053-510-3            Sequence 3, Appli
     43   265.5   10.2   1629  14  US-10-286-175-5            Sequence 5, Appli
c    44     257    9.8    359  12  US-10-084-843-192          Sequence 192, App
c    45     257    9.8    359  12  US-10-193-002-187          Sequence 187, App

                                 ALIGNMENTS

RESULT 1
US-10-006-852-1
; Sequence 1, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1509
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1509)
;   OTHER INFORMATION:
US-10-006-852-1
```

```
lignment Scores:
red. No.:              8.08e-295      Length:        1509
core:                  2615.00        Matches:       502
ercent Similarity:     100.00%        Conservative:  0
est Local Similarity:  100.00%        Mismatches:    0
uery Match:            100.00%        Indels:        0
B:                     14             Gaps:          0

S-10-006-852-2 (1-502) x US-10-006-852-1 (1-1509)

y          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

y         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

y         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

y         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

y         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

y        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

y        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

y        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

y        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

y        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

y        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

y        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

y        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

y        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

y        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

y        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy       321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy       341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy       361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy       381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy       401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy       421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy       441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Qy       461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy       481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

Qy       501 IleCys 502
             ||||||
Db      1501 ATCTGC 1506

RESULT 2
US-09-938-842A-1810
; Sequence 1810, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1810
;   LENGTH: 1509
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-1810

Alignment Scores:
Pred. No.:             6.94e-294      Length:        1509
Score:                 2607.00        Matches:       501
Percent Similarity:    99.80%         Conservative:  0
Best Local Similarity: 99.80%         Mismatches:    1
```

```
ery Match:              99.69%         Indels:        0
:                       10             Gaps:          0

-10-006-852-2 (1-502) x US-09-938-842A-1810 (1-1509)

             1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
             1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

            21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

            41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

            61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

            81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

           101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

y          121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

y          141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

y          161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

y          181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

y          201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b          601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

y          221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

y          241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

y          261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

y          281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

y          301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

y          321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy         341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy         361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy         381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy         401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy         421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy         441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Qy         461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy         481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

Qy         501 IleCys 502
               ||||||
Db        1501 ATCTGC 1506

RESULT 3
US-10-006-852-15
; Sequence 15, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 1785
;   TYPE: DNA
;   ORGANISM: Petunia x hybrida
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (72)..(1571)
;   OTHER INFORMATION:
US-10-006-852-15

Alignment Scores:
Pred. No.:              4.19e-252      Length:        1785
Score:                  2250.00        Matches:       427
Percent Similarity:     92.64%         Conservative:  39
Best Local Similarity:  84.89%         Mismatches:    33
Query Match:            86.04%         Indels:        4
DB:                     14             Gaps:          2

US-10-006-852-2 (1-502) x US-10-006-852-15 (1-1785)

Qy           1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
               ||||||||||||      ||||||:::||||||||||||:::||||||||||||||||||
```

```
       72 ATGGTTCTATCAAAGACAGTGTCGCAGAGCGATGTGTCCATTCACTCCACGTTTGCTTCT 131

       21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
          |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
      132 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGATAATTCGATACCAAAAGAA 191

       41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      192 GCAGCATATCAGATCATAAATGATGAACTGATGTTAGATGGAAACCCAAGGCTGAACTTG 251

       61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
          |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
      252 GCTTCTTTTGTTACAACATGGATGGAACCAGAGTGTGATAAGTTGATGATGGACTCTATT 311

       81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      312 AACAAGAACTATGTTGATATGGATGAATATCCTGTTACCACTGAGCTTCAGAATCGATGT 371

      101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
          |||||||||||||||||||||||||||||||||||||||:::   |||||||||||||||
      372 GTAAACATGATAGCTCATTTGTTTAATGCACCACTTGAAGATGGAGAAACTGCAGTTGGA 431

      121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      432 GTTGGAACTGTTGGATCCTCTGAAGCCATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 491

      141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
          ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
      492 TGGCAGAACAAAATGAAAGCCCAAGGCAAACCCTGTGACAAGCCCAACATTGTTACTGGT 551

      161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      552 GCAAATGTCCAGGTGTGCTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 611

      181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
          ||||||||||||||||||||||||||||||||||||::::::||||||:::|||||||||
      612 GTAAAGCTTAGTGAAGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAGATGGTGGAT 671

      201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
          |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
      672 GAAAACACCATTTGTGTAGCTGCTATCTTAGGTTCCACCCTCAATGGAGAATTTGAAGAC 731

      221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
          ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
      732 GTTAAGCGCTTGAATGATCTCTTGGTCGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 791

y     241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
          ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
b     792 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCGTTCATTTACCCAGAGCTTGAG 851

y     261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     852 TGGGACTTTAGATTGCCATTAGTGAAGAGCATTAATGTAAGTGGTCACAAATATGGTCTT 911

y     281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
          ||||||||||||||||||||||||:::||||||||||||:::|||||||||:::||||||
b     912 GTCTATGCTGGTATTGGTTGGGTCGTTTGGAGGAACAAGGATGATTTGCCTGATGAACTT 971

y     301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     972 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTTTCTAAA 1031

y     321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
          ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
b    1032 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTATGAGGGTTAC 1091

y     341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
          :::||||||||||||||||||:::||||||      ||||||||||||||||||||||||
b    1092 AAGAATGTGATGGAGAATTGTCAAGAAAATGCATCGGTACTAAGAGAAGGGCTAGAAAAG 1151

y     361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
          |||   ||||||||||||:::||||||:::   ||||||||||||||||||||||||||
b    1152 ACAGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTACCTTTAGTAGCATTCTCTCTT 1211
```

```
Qy        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||:::      |||   |||||||||||||||:::   |||||||||:::||||||
Db       1212 AAAGACAACAGGCAACACAACGAGTTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1271

Qy        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1272 ATTGTTCCTGCATATACTATGCCACCAAACGCACAACACATTACAGTTCTCAGAGTTGTG 1331

Qy        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db       1332 ATCAGAGAAGATTTCTCCCGTACGCTTGCAGAACGACTGGTAAGAGACATCGAAAAAGTC 1391

Qy        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              :::   |||||||||   ||||||:::||||||      |||:::::::::   :::|||
Db       1392 CTTCATGAACTTGACACACTCCCTGCACGTGTCAATGCTAAGCTCGCTGTGGCCGAGGAG 1451

Qy        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
              ::::::   :::|||         :::            ||||||:::|||   :::
Db       1452 CAGGCGGCTGCGAATGGCAGCGAGGTGCAT---------AAGAAAACAGATAGCGAAGTG 1502

Qy        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
              |||   ::::::||||||   |||||||||||||||   :::   :::|||||||||:::
Db       1503 CAGTTGGAGATGATAACTGCATGGAAGAAGTTTGTTGAAGAAAAGAAGAAGAAGACTAAT 1562

Qy        500 GlyIleCys 502
                 :::|||
Db       1563 CGAGTTTGT 1571

RESULT 4
US-10-006-852-11
; Sequence 11, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 1705
;   TYPE: DNA
;   ORGANISM: Nicotiana tabacum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (71)..(1558)
;   OTHER INFORMATION:
US-10-006-852-11

Alignment Scores:
Pred. No.:              5.72e-251      Length:        1705
Score:                  2240.00        Matches:       428
Percent Similarity:     92.05%         Conservative:  35
Best Local Similarity:  85.09%         Mismatches:    32
Query Match:            85.66%         Indels:        8
DB:                     14             Gaps:          3

US-10-006-852-2 (1-502) x US-10-006-852-11 (1-1705)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||         ||||||||||||||||||:::||||||||||||||||||
Db         71 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 130

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 CGATATGTTCGTACTTCTCTTCCGAGGTTTAAGATGCCAGAGAATTCGATACCAAAGGAA 190

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           191 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGACTAAATTTA 250

            61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
               ||||||||||||||||||||||||||||||||||||:::||||||:::|||   ||||||
           251 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAACAAACTGATGATGGATTCCATT 310

            81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           311 AACAAGAATTACGTTGACATGGATGAATACCCTGTAACCACTGAACTTCAGAATCGATGT 370

           101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
               ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
           371 GTAAACATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 430

           121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           431 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 490

           141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
               ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
           491 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTGTGACAAGCCCAATATTGTCACTGGT 550

           161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           551 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 610

y          181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
               ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
b          611 GTAAAGTTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 670

y          201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b          671 GAGAACACAATTTGTGTAGCTGCTATCTTGGGTTCCACACTCAATGGTGAATTTGAAGAT 730

y          221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
               ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
b          731 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 790

y          241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          791 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 850

y          261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          851 TGGGACTTTAGATTGCCATTGGTGAAGAGTATAAACGTGAGTGGTCACAAATATGGTCTT 910

y          281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
               |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
b          911 GTTTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTACCTGACGAACTT 970

y          301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          971 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1030

y          321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
               ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
b         1031 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1090

y          341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
               :::||||||||||||||||||:::||||||      ||||||||||||||||||||||||
b         1091 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1150

y          361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
               :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
b         1151 AGTGGAAGATTCAACATAATATCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1210

y          381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
               ||||||:::|||   |||   |||||||||||||||:::   |||||||||:::||||||
b         1211 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1270

y          401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
               |||:::|||||||||||||||||||||||||||||||||:::||||||||||||||||||

Db        1271 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1330

Qy         421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1331 ATTAGAGAAGATTTCTCCCGTACACTCGCCGAGCGACTGGTAATAGACATTGAAAAAGTC 1390

Qy         441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
               :::   |||||||||   ||||||:::||||||      |||:::::::::
Db        1391 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTG--------- 1441

Qy         461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                  :::|||:::|||         :::            ||||||:::|||   :::
Db        1442 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1489

Qy         481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
               |||   :::|||   |||   |||||||||||||||||||||   :::|||||||||:::
Db        1490 CAGCTTGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAC 1549

Qy         500 GlyIleCys 502
               |||:::|||
Db        1550 GGAGTTTGT 1558

RESULT 5
US-10-005-602-1
; Sequence 1, Application US/10005602
; Publication No. US20030110530A1
; GENERAL INFORMATION:
;  APPLICANT: Barry Shelp
;  APPLICANT:  Alan Bown
;  TITLE OF INVENTION: TRANSGENIC PLANTS HAVING REDUCED
;  TITLE OF INVENTION:  SUSCEPTIBILITY TO INVERTEBRATE PESTS
;  FILE REFERENCE: P84US3
;  CURRENT APPLICATION NUMBER: US/10/005,602
;  CURRENT FILING DATE:  2002-12-31
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1745
;   TYPE: DNA
;   ORGANISM: tobacco plant
US-10-005-602-1

Alignment Scores:
Pred. No.:               3.89e-250      Length:         1745
Score:                   2233.00        Matches:        427
Percent Similarity:      91.85%         Conservative:   35
Best Local Similarity:   84.89%         Mismatches:     33
Query Match:             85.39%         Indels:         8
DB:                      14             Gaps:           3

US-10-006-852-2 (1-502) x US-10-005-602-1 (1-1745)

Qy           1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
               ||||||||||||         ||||||||||||||||||:::||||||||||||||||||
Db          31 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 90

Qy          21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          91 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAGATGCCAGAGAATTCAATACCAAAGGAA 150

Qy          41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         151 GCAGCATATCAGATTATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 210

Qy          61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
               ||||||||||||||||||||||||||||||||||||:::   |||:::|||   ||||||
Db         211 GCATCTTTCGTTACAACATGGATGGAGCCAGAATGTAATACGTTAATGATGGATTCCATT 270

Qy          81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         271 AACAAGAACTACGTTGACATGGATGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 330

Qy         101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
```

```
               ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
b         331 GTAAATATGATAGCTCATTTGTTTAATGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 390

y         121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         391 GTTGGAACTGTTGGATCCTCTGAAGCTATTATGCTTGCTGGATTAGCCTTTAAAAGAAAA 450

y         141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
b         451 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTTTGATAAGCCCAATATCGTCACCGGT 510

y         161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         511 GCTAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 570

y         181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
b         571 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 630

y         201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b         631 GAGAATACCATTTGTGTTGCTGCTATCTTAGGTTCAACACTCAATGGTGAATTTGAAGAT 690

y         221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
b         691 GTTAAGCGTTTGAATGACCTTTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 750

y         241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         751 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 810

y         261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         811 TGGGACTTTAGATTGCCATTGGTGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 870

y         281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
b         871 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 930

y         301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         931 ATTTTCCACATCAATTACCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 990

y         321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
b         991 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1050

y         341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              :::||||||||||||||||||:::||||||      |||||||||||||||:::||||||
b        1051 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTATTAAGAGAAGGAATTGAAAAA 1110

y         361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
b        1111 AGTGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTTCCCTTAGTAGCATTTTCTCTT 1170

y         381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||:::|||   |||   |||||||||||||||:::   |||||||||:::||||||
b        1171 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1230

y         401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
b        1231 ATTGTTCCTGCATATACTATGCCACCAAATGCTCAACATGTTACAGTTCTCAGAGTTGTC 1290

y         421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        1291 ATTAGAGAAGATTTCTCCCGCACACTAGCGGAGCGACTGGTAATAGACATTGAAAAAGTC 1350

y         441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              :::   |||||||||   ||||||:::||||||      |||:::::::::
b        1351 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTG--------- 1401

y         461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                 :::|||:::|||         :::            ||||||:::|||   :::
```

```
Db       1402 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1449

Qy        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
              |||   :::|||   |||   |||||||||||||||||||||   :::|||||||||:::
Db       1450 CAGCTAGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAT 1509

Qy        500 GlyIleCys 502
              |||:::|||
Db       1510 GGAGTTTGT 1518

RESULT 6
US-09-887-576-820
; Sequence 820, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
;  APPLICANT: Budworth, P.
;  APPLICANT:  Brown, D.
;  APPLICANT:  Chang, H.
;  APPLICANT:  Zhu, T.
;  APPLICANT:  Han, B.
;  APPLICANT:  Wang, X.
;  APPLICANT:  Cooper, Bret
;  TITLE OF INVENTION: Promoters for regulation of plant expression
;  FILE REFERENCE: 1360.001US1
;  CURRENT APPLICATION NUMBER: US/09/887,576
;  CURRENT FILING DATE:  2001-06-25
;  PRIOR APPLICATION NUMBER: US 60/213,848
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/214,087
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/258,692
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 875
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 820
;   LENGTH: 1479
;   TYPE: DNA
;   ORGANISM: Oryza sativa
US-09-887-576-820

Alignment Scores:
Pred. No.:               2.13e-246     Length:          1479
Score:                   2200.00       Matches:         420
Percent Similarity:      90.64%        Conservative:    35
Best Local Similarity:   83.67%        Mismatches:      37
Query Match:             84.13%        Indels:          10
DB:                      10            Gaps:            2

US-10-006-852-2 (1-502) x US-09-887-576-820 (1-1479)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||   ||||||||||||||||||:::||||||||||||||||||||||||
Db          1 ATGGTGCTCTCCAAGGCCGTCTCCGAGAGTGACATGTCCGTGCACTCCACCTTCGCCTCC 60

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||   ||||||||||||::::::|||||||||||||||||||||||||||
Db         61 CGCTACGTCCGCGCCTCCCTCCCAAGGTACCGGATGCCGGAGAACTCGATCCCGAAGGAG 120

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGCGTACCAGATCATCAACGACGAGCTGATGCTGGACGGCAACCCGCGGCTGAACCTG 180

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||::::::|||
Db        181 GCGTCGTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATCATGGCCGCCATC 240

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAGAACTACGTCGACATGGACGAGTACCCCGTCACCACCGAGCTCCAGAACCGGTGC 300

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||||||||||||||||||||:::|||||||||   |||   |||||||||||||||
Db        301 GTGAACATGATCGCGCACCTGTTCCACGCGCCGCTCGGGGAGGACGAGACGGCGGTGGGC 360
```

```
y       121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
b       361 GTGGGCACGGTGGGTTCGTCGGAGGCCATCATGCTGGCCGGGCTGGCCTTCAAGCGGCGG 420

y       141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
            |||||||||||||||||||||||||||||||||   |||||||||||||||:::||||||
b       421 TGGCAGAACAAGCGCAAGGCCGAGGGGAAGCCGTTCGACAAGCCCAACATCATCACCGGC 480

y       161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       481 GCCAACGTGCAGGTGTGCTGGGAGAAGTTCGCCCGCTACTTCGAGGTGGAGCTCAAGGAG 540

y       181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
            |||||||||   :::|||||||||||||||||||||::::::|||||||||||||||:::
b       541 GTGAAGCTCCGCGACGGCTACTACGTCATGGACCCCGAGAAGGCCGTCGACATGGTCAAC 600

y       201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
            |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b       601 GAGAACACCATCTGCGTCGCCGCCATCCTCGGCTCCACCCTCAACGGCGAGTTCGAGGAC 660

y       221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
            ||||||||||||||||||||||||   :::|||||||||||||||||||||:::||||||
b       661 GTCAAGCTACTCAACGACCTCCTCGACAAGAAGAACAAGGAGACTGGGTGGGAGACGCCG 720

y       241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       721 ATCCACGTGGACGCGGCGAGCGGCGGGTTCATCGCGCCGTTCCTGTACCCGGAGCTGGAG 780

y       261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
            ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
b       781 TGGGACTTCCGGCTGCCGTGGGTGAAGAGCATCAACGTGAGCGGTCACAAGTACGGGCTC 840

y       281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
            |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b       841 GTCTACGCCGGCATCGGCTGGTGCATCTGGCGCAACAAGGAGGACCTGCCCGAGGAGCTC 900

y       301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
            ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
b       901 ATCTTCCACATCAACTACCTCGGCACCGACCAGCCAACCTTCACCCTCAACTTCTCCAAG 960

y       321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
            ||||||||||||||||||||||||||||||||||||||||||   |||   |||||||||
b       961 GGCTCCAGCCAGGTCATCGCCCAGTACTACCAGCTCATCCGCCACGGCTTCGAGGGGTAC 1020

y       341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
            ||||||:::||||||||||||   ||||||   :::||||||:::|||||||||   |||
b      1021 AGGAACATCATGGAGAACTGCCACGAGAACGCGATGGTGCTGAAGGAAGGGCTGGTGAAG 1080

y       361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
            |||   ||||||:::|||||||||||||||||||||||||||||||||||||||||||||
b      1081 ACGGGGAGGTTCGACATCGTGTCCAAGGACGAAGGGGTGCCGCTGGTGGCGTTCTCGCTC 1140

y       381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
            ||||||   |||   |||   ||||||||||||||||||||||||||||||:::||||||
b      1141 AAGGACCGGAGCCGGCACGACGAGTTCGAGATCTCCGACATGCTGCGCCGCTTCGGCTGG 1200

y       401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
            |||||||||||||||||||||||||||:::|||||||||:::||||||||||||||||||
b      1201 ATCGTGCCGGCGTACACCATGCCGCCCGACGCCCAGCACGTCACGGTGCTCCGCGTGGTC 1260

y       421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
            |||||||||:::||||||||||||||||||||||||||||||:::|||||||||||||||
b      1261 ATCCGGGAGGAGTTCAGCCGCACCCTCGCCGAGCGCCTCGTCCTCGACATCGAGAAGGTG 1320

y       441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
            |||   :::||||||   ||||||||||||::::::      :::
b      1321 ATGTACCAGCTCGACGCGCTCCCCTCCAGGCTCATGCCCCCCGTGCCGCCGGCGCCG--- 1377

y       461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                                    |||:::|||      |||||||||:::::::::
b      1378 ------------------------CTGCTGGTGGTCGCCAAGAAGTCGGAGCTCGAGACG 1413
```

```
Qy      481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
            ||||||   :::         |||||||||||||||         |||:::|||:::|||
Db     1414 CAGCGGTCGGTGACGGAGGCGTGGAAGAAGTTCGTGCTC---GCCAAGAGGACCAACGGC 1470

Qy      501 IleCys 502
            :::|||
Db     1471 GTCTGC 1476

RESULT 7
US-10-006-852-13
; Sequence 13, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 1771
;   TYPE: DNA
;   ORGANISM: Nicotiana tabacum
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (67)..(1554)
;   OTHER INFORMATION:
US-10-006-852-13

Alignment Scores:
Pred. No.:              3.7e-246      Length:         1771
Score:                  2199.00       Matches:        423
Percent Similarity:     90.85%        Conservative:   34
Best Local Similarity:  84.10%        Mismatches:     38
Query Match:            84.09%        Indels:         8
DB:                     14            Gaps:           3

US-10-006-852-2 (1-502) x US-10-006-852-13 (1-1771)

Qy        1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
            ||||||||||||         |||||||||||||||||||||||||||||||||||||||
Db       67 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCGTTCACTCCACTTTCGCCTCC 126

Qy       21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      127 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGAGAATTCAATACCAAAGGAA 186

Qy       41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      187 GCAGCATATCAGATTATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 246

Qy       61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
            ||||||||||||||||||||||||||||||||||||:::   |||:::|||   ||||||
Db      247 GCATCTTTCGTTACAACATGGATGGAGCCAGAATGTAATACGTTAATGATGGATTCCATT 306

Qy       81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      307 AACAAGAACTACGTTGACATGGATGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 366

Qy      101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
            ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
Db      367 GTAAATATGATAGCTCATTTGTTTAATGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 426

Qy      121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      427 GTTGGAACTGTTGGATCCTCTGAAGCTATTATGCTTGCTGGATTAGCCTTTAAGAGAAAA 486

Qy      141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
            ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
```

```
b        487 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTTTGATAAGCCCAATATTGTCACCGGT 546

y        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        547 GCTAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 606

y        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
             ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
b        607 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 666

y        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
             |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b        667 GAGAATACCATTTGTGTTGCTGCTATCTTAGGTTCAACACTCAATGGTGAATTTGAAGAT 726

y        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
             ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
b        727 GTTAAGCGTTTGAATGACCTTTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 786

y        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        787 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 846

y        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b        847 TGGGACTTTAGATTGCCATTGGAGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 906

y        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
b        907 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 966

y        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        967 ATTTTCCACATCAATTACCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1026

y        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
b       1027 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1086

y        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             :::||||||||||||||||||:::||||||      |||||||||||||||:::||||||
b       1087 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTATTAAGAGAAGGAATTGAAAAA 1146

y        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
             :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
b       1147 AGTGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTTCCCTTAGTAGCATTTTCTCTT 1206

y        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
             ||||||:::|||   |||   |||||||||||||||:::   |||||||||:::||||||
b       1207 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1266

y        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
             ||||||   ||||||||||||||||||||||||||||||:::||||||||||||||||||
b       1267 ATTGTTCTGGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1326

y        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       1327 ATTAGAGAAGATTTCTCCCGCACACTAGCGGAGCGACTGGTAATAGACATTGAAAAAGTC 1386

y        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
                      :::|||   ||||||:::||||||      |||:::::::::
b       1387 TTCCACGGAGTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTG--------- 1437

y        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                :::|||:::|||         :::            ||||||:::|||   :::
b       1438 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1485

y        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
             |||   :::|||   |||   |||   |||||||||||||||   :::|||||||||:::
b       1486 CAGCTAGAGATTACTACTGCATGGTTGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAT 1545

y        500 GlyIleCys 502
             |||:::|||
b       1546 GGAGTTTGT 1554
```

```
RESULT 8
US-09-938-842A-937
; Sequence 937, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 937
;   LENGTH: 1482
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-937

Alignment Scores:
Pred. No.:                7.22e-239       Length:        1482
Score:                    2135.50         Matches:       411
Percent Similarity:       88.78%          Conservative:  32
Best Local Similarity:    82.36%          Mismatches:    45
Query Match:              81.66%          Indels:        11
DB:                       10              Gaps:          2

US-10-006-852-2 (1-502) x US-09-938-842A-937 (1-1482)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||      |||||||||||||||||||||:::||||||||||||||||||
Db          1 ATGGTTTTGTCTAAGACAGTTTCCGAATCTGATGTCTCAATCCATTCAACTTTTGCTTCT 60

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||   |||||||||||||||:::|||||||||||||||||||||||||||
Db         61 CGTTACGTCCGCAACTCTCTTCCACGATTCGAAATGCCTGAGAACTCAATCCCAAAAGAA 120

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCAGCTTACCAAATCATCAACGACGAGCTAATGCTCGATGGTAACCCAAGGCTGAACCTA 180

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
Db        181 GCTTCCTTCGTGACCACATGGATGGAGCCAGAATGTGACAAGCTCATGATGGAGTCCATC 240

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAGAACTACGTCGACATGGACGAGTACCCTGTCACCACTGAGCTTCAGAACCGATGT 300

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              |||||||||||||||   ||||||||||||||||||   :::   |||   |||||||||
Db        301 GTTAACATGATAGCACGTCTCTTCAACGCGCCGCTTGGTGACGGTGAAGCTGCCGTTGGT 360

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        361 GTTGGCACCGTCGGATCGTCGGAGGCGATTATGTTGGCCGGTTTGGCTTTTAAGAGACAA 420

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              |||||||||||||||||||||:::|||   |||   ||||||||||||||||||||||||
Db        421 TGGCAGAATAAGCGTAAGGCCCAAGGGCTTCCTTATGATAAGCCCAATATCGTAACCGGT 480

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTAATGTCCAGGTTTGCTGGGAGAAATTCGCAAGGTATTTCGAAGTGGAGCTTAAGGAA 540
```

```
/         181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              |||   |||   |||   ||||||||||||||||||   :::||||||:::|||||||||
)         541 GTGAACCTAAGAGAAGACTATTACGTGATGGACCCTGTAAAGGCGGTCGAAATGGTAGAC 600

/         201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||   |||||||||||||||
)         601 GAAAACACAATTTGTGTCGCTGCCATCCTCGGTTCAACGTTAACCGGTGAATTCGAAGAC 660

/         221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
)         661 GTTAAGCTCCTCAACGACCTCCTTGTCGAGAAAAACAAGCAAACCGGATGGGACACGCCA 720

/         241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)         721 ATACACGTGGACGCAGCGAGTGGTGGGTTTATTGCTCCGTTCTTGTATCCGGAGCTGGAG 780

/         261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)         781 TGGGATTTCCGGCTACCGTTGGTTAAGAGTATTAATGTGAGTGGTCACAAATACGGTTTG 840

/         281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              ||||||||||||||||||||||||:::||||||   |||   |||||||||:::||||||
)         841 GTTTACGCCGGTATTGGTTGGGTTGTATGGAGAACCAAAACCGATTTGCCTGATGAACTT 900

y         301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)         901 ATCTTCCATATCAATTATCTTGGCGCTGATCAACCAACCTTTACACTCAACTTCTCCAAA 960

y         321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
b         961 GGTTCAAGTCAAGTGATTGCTCAGTACTACCAGCTGATTCGTCTTGGATTCGAGGGTTAT 1020

y         341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              ||||||||||||:::||||||||||||||||||:::|||||||||:::||||||||||||
b        1021 CGCAATGTGATGGATAATTGTCGGGAAAACATGATGGTACTAAGACAAGGATTAGAGAAA 1080

y         361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              |||   ||||||   ||||||||||||:::   |||||||||||||||||||||||||||
b        1081 ACGGGACGTTTTAAAATCGTCTCCAAAGAAAACGGTGTTCCGTTAGTGGCGTTTTCTCTC 1140

y         381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||||||||   |||   |||||||||::::::      |||||||||:::||||||
b        1141 AAAGATAGTAGCCGCCACAACGAGTTCGAGGTGGCCCATACACTCCGTCGCTTCGGCTGG 1200

y         401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              ||||||||||||||||||||||||   :::|||||||||:::||||||||||||||||||
b        1201 ATCGTTCCGGCCTACACGATGCCTGCGGATGCGCAGCATGTCACTGTCCTTCGAGTTGTT 1260

y         421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
              ||||||||||||||||||||||||||||||||||||||||||   |||   |||||||||
b        1261 ATCCGAGAAGATTTCTCTCGAACCTTAGCCGAGAGATTGGTAGCTGATTTCGAGAAGGTT 1320

y         441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              :::   |||||||||   ||||||:::||||||      |||::::::   |||:::
b        1321 CTACACGAGCTCGATACGCTTCCGGCGAGGGTTCACGCCAAGATGGCTAATGGAAAAGTT 1380

y         461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                                                  |||||||||:::      :::
b        1381 AACGGT------------------------------GTTAAGAAGACGCCAGAGGAGACG 1410

y         481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSer 499
              ||||||::::::         |||||||||   :::   :::   |||||||||:::
)b       1411 CAGAGAGAAGTCACGGCCTACTGGAAGAAGTTGTTG---GAGACTAAGAAGACCAAC 1464

RESULT 9
JS-10-006-852-3
 Sequence 3, Application US/10006852
 Publication No. US20030046732A1
 GENERAL INFORMATION:
  APPLICANT: Kinnersely, Alan M.
  APPLICANT:  Turano, Frank J.
  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1665
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (17)..(1498)
;   OTHER INFORMATION:
US-10-006-852-3

Alignment Scores:
Pred. No.:                3.8e-238      Length:         1665
Score:                    2130.00       Matches:        404
Percent Similarity:       89.29%        Conservative:   46
Best Local Similarity:    80.16%        Mismatches:     42
Query Match:              81.45%        Indels:         12
DB:                       14            Gaps:           3

US-10-006-852-2 (1-502) x US-10-006-852-3 (1-1665)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              |||||||||:::         :::   :::|||   ||||||   :::   |||   |||
Db         17 ATGGTTTTGACAAAAACCGCAACG---AATGATGAATCTGTCTGCACCATGTTCGGATCT 73

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              |||||||||||||||:::||||||::::::::::::   ||||||||||||||||||:::
Db         74 CGCTATGTTCGCACTACACTTCCCAAGTATGAGATTGGTGAGAATTCGATACCGAAAGAC 133

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        134 GCTGCATATCAGATCATAAAAGATGAGCTGATGCTTGATGGTAACCCGAGGCTTAACCTA 193

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        194 GCTTCGTTTGTGACTACATGGATGGAACCAGAGTGTGACAAACTCATCATGGACTCTATC 253

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        254 AACAAGAACTACGTTGATATGGATGAGTACCCTGTCACAACTGAGCTCCAGAACCGATGT 313

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||:::||||||   ||||||||||||||||||||||||:::|||||||||||||||
Db        314 GTAAACATTATAGCTCGACTGTTCAATGCGCCACTCGAGGAATCTGAGACGGCGGTGGGA 373

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 GTAGGGACAGTTGGTTCTTCAGAAGCCATCATGTTAGCCGGATTGGCCTTCAAAAGAAAA 433

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db        434 TGGCAGAACAAACGCAAGGCTGAGGGTAAACCCTATGACAAACCCAACATTGTCACTGGA 493

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        494 GCCAATGTTCAAGTTTGCTGGGAGAAATTCGCTCGGTACTTCGAGGTGGAGCTAAAGGAA 553

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              |||   ||||||||||||||||||||||||||||||   :::|||   :::|||||||||
Db        554 GTAAACCTAAGTGAAGGTTACTACGTGATGGATCCAGACAAAGCAGCAGAAATGGTAGAC 613

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        614 GAGAACACAATCTGTGTCGCAGCCATATTGGGATCCACACTCAACGGTGAGTTCGAAGAC 673

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   ||||||||||||||||||:::||||||:::||||||||||||:::||||||
```

```
          674 GTGAAACGTCTCAATGACTTGCTAGTCAAGAAAAACGAGGAGACTGGTTGGAACACACCG 733

          241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
          734 ATCCACGTGGATGCAGCAAGTGGAGGGTTCATAGCTCCGTTTATCTATCCTGAATTAGAA 793

          261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          794 TGGGACTTTAGACTTCCTTTGGTTAAGAGTATCAACGTGAGTGGTCACAAGTATGGACTG 853

          281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              ||||||||||||||||||||||||:::||||||      |||||||||||||||||||||
          854 GTCTATGCTGGTATTGGTTGGGTCGTGTGGAGGGCAGCAGAGGATTTGCCTGAAGAGCTT 913

          301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          914 ATCTTTCATATTAATTATCTTGGTGCTGATCAACCCACTTTCACTCTCAATTTCTCCAAG 973

          321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||:::|||||||||||||||||||||||||||||||||   |||||||||
          974 GGATCGAGCCAAATTATTGCTCAATACTACCAGCTCATTCGTCTTGGATTCGAGGGGTAC 1033

          341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              :::||||||||||||||||||   |||||||||:::||||||:::||||||:::||||||
         1034 AAAAATGTGATGGAGAATTGCATAGAGAACATGGTGGTTCTCAAAGAAGGGATAGAGAAA 1093

          361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              ||||||||||||||||||||||||||||||:::|||||||||:::|||||||||||||||
         1094 ACAGAGCGTTTCAACATAGTCTCAAAGGACCAAGGAGTGCCAGTCGTAGCCTTCTCTCTC 1153

          381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||   |||   |||   |||||||||||||||:::||||||||||||:::||||||
         1154 AAGGACCATAGTTTCCACAACGAGTTCGAGATCTCTGAGATGCTACGTCGTTTTGGCTGG 1213

y         401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              ||||||||||||||||||||||||   :::||||||||||||||||||||||||||||||
b        1214 ATCGTCCCAGCTTACACTATGCCTGCCGATGCACAGCACATCACGGTTCTGCGTGTTGTC 1273

y         421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
              ||||||||||||||||||||||||||||||||||||||||||   ||||||   ||||||
b        1274 ATCAGGGAAGATTTCTCAAGAACACTCGCGGAGAGACTTGTTGCTGATATTTCGAAGGTG 1333

y         441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              :::   |||||||||   |||||||||::::::      |||:::   :::
b        1334 CTTCATGAGCTAGATACCTTGCCTTCCAAGATATCTAAGAAGATGGGAATA--------- 1384

y         461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                                :::   :::         ||||||:::      ::::::|||
b        1385 ------------------GAAGGGATCGCGGAAAATGTAAAGGAGAAGAAGATGGAGAAG 1426

y         481 Gln------ArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThr 498
              :::         ::::::|||   ||||||:::|||||||||   :::|||||||||
b        1427 GAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAGTTTGTGAAGGAGAGGAAGAAGATG 1486

y         499 SerGlyIleCys 502
              :::|||:::|||
b        1487 AATGGTGTGTGC 1498

ESULT 10
S-10-167-547C-1
 Sequence 1, Application US/10167547C
 Publication No. US20030170653A1
 GENERAL INFORMATION:
  APPLICANT: E.I. du Pont de Nemours and Company
  APPLICANT:  Damude, Howard G.
  TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
  TITLE OF INVENTION:  Butyrolactone and its Intermediates
  FILE REFERENCE: CL1804 US NA
  CURRENT APPLICATION NUMBER: US/10/167,547C
  CURRENT FILING DATE:  2003-03-17
  PRIOR APPLICATION NUMBER: 60/297198
  PRIOR FILING DATE: 2001-06-08
  NUMBER OF SEQ ID NOS: 67
```

```
;  SOFTWARE: Microsoft Office 07
; SEQ ID NO 1
;   LENGTH: 1497
;   TYPE: DNA
;   ORGANISM: Alstroemeria
US-10-167-547C-1

Alignment Scores:
Pred. No.:              1.44e-231      Length:        1497
Score:                  2073.00        Matches:       391
Percent Similarity:     87.92%         Conservative:  53
Best Local Similarity:  77.43%         Mismatches:    51
Query Match:            79.27%         Indels:        10
DB:                     12             Gaps:          2

US-10-006-852-2 (1-502) x US-10-167-547C-1 (1-1497)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||   |||||||||::::::         ||||||   ||||||||||||
Db          1 ATGGTTCTCTCCAGCGCCGTCTCCGACAACACCGGCCCAGTCCACTGCACCTTCGCGTCG 60

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||   :::      ||||||:::||||||||||||||||||||||||:::
Db         61 CGCTACGTCCGCGATGCGCCTGCCCGGTTCAGGATGCCGGAGAATTCGATACCCAAGGAC 120

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
                 ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        121 ACGGCGTACCAGATCGTCAACGACGAGCTGATGCTGGACGGGAACCCCAGGCTGAACCTG 180

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||:::|||::::::::::::
Db        181 GCGTCGTTCGTCACGACCTGGATGGAGCCGGAGTGCGATCGCCTCATGATCGCCGCCGCC 240

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAGAACTATGTCGACATGGACGAGTACCCGGTCACAACCGAGCTACAGAATCGCTGT 300

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||:::||||||||||||||||||||||||:::   :::   |||||||||||||||
Db        301 GTTAATATAATAGCCCACCTCTTCAATGCGCCTATTGGGGATGAGGAAACAGCAGTAGGA 360

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTTGGGACAGTGGGGTCATCAGAAGCAATAATGCTTGCAGGCTTGGCATTCAAGAGAAAG 420

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||:::||||||||||||||||||   ||||||||||||||||||||||||
Db        421 TGGCAGAACAAAAAAAAAGCAGAGGGGAAGCCATATGACAAGCCCAATATTGTCACCGGT 480

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCAAACGTTCAGGTTTGCTGGGAGAAATTCGCTAGGTATTTTGAAGTTGAACTGAAAGAA 540

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              |||||||||   ||||||||||||:::|||||||||::::::||||||:::|||||||||
Db        541 GTGAAGCTGAGGGAGGGTTATTACATCATGGACCCAGAGAAGGCTGTGGAAATGGTGGAT 600

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||   |||||||||||||||
Db        601 GAGAATACCATATGTGTTGCTGCTATCTTGGGCTCAACCCTTACTGGAGAGTTCGAAGAT 660

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              |||||||||||||||   ||||||   |||||||||||||||||||||||||||||||||
Db        661 GTTAAACTACTGAACAAACTTCTTGAAGAGAAGAACAAGGAAACTGGGTGGGACACACCC 720

Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATTCATGTTGATGCTGCTAGTGGTGGATTCATTGCTCCTTTTCTATACCCAGAACTGGAA 780

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGATTTCCGATTACCACTGGTGAAGAGTATTAATGTCAGCGGACACAAATATGGCCTT 840
```

```
y        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             ||||||||||||:::|||||||||:::|||||||||||||||||||||||||||||||||
b        841 GTTTATGCAGGTGTGGGTTGGGTTGTCTGGAGGAACAAAGAAGATCTTCCTGAAGAGCTC 900

y        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        901 ATTTTCCATATAAACTACCTTGGGGCAGATCAGCCTACTTTCACCCTCAATTTCTCTAAA 960

y        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||:::||||||||||||||||||   ||||||||||||   :::||||||
b        961 GGTTCAAGCCAGATAATTGCTCAATATTATCAATTCATTCGCCTTGGTTTTCAGGGGTAT 1020

y        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             :::|||:::||||||||||||   ||||||      :::|||||||||||||||::::::
b       1021 AAGAACATAATGGAAAACTGCATGGAGAACACAAGAATACTGAGAGAAGGTCTGCAGGAG 1080

y        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
             |||   ||||||   |||||||||||||||   |||||||||||||||||||||:::|||
b       1081 ACGGGCCGTTTCGAGATAGTCTCCAAAGATATTGGGGTGCCTCTTGTTGCATTTGCTCTC 1140

y        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
             ||||||||||||   :::|||   |||||||||:::|||   :::||||||:::||||||
b       1141 AAGGACAGCAGCCAGTACACTGTCTTTGAGATAGCGGACGCCATGAGAAGGTTCGGATGG 1200

y        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
             |||:::||||||||||||||||||   :::|||:::||||||   |||||||||||||||
b       1201 ATCATTCCTGCATACACCATGCCAAAGGACGCGGAGCACATAGCTGTCCTCCGTGTGGTT 1260

y        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
             |||||||||||||||||||||:::||||||||||||||||||   |||::::::||||||
b       1261 ATCAGGGAGGATTTCAGCAGGAGCCTTGCTGAGCGCCTAGTTAATGACATGAAGAAGGTG 1320

y        441 MetArgGluLeuAspGluLeuProSerArg---------ValIleHisLysIleSerLeu 457
             :::   |||||||||   ||||||||||||         :::   |||      ::::::
b       1321 CTGGCTGAGCTGGACGTACTTCCCAGTCGCATCACCACCATTGCCCATGTTACGGCTGTG 1380

y        458 GlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAsp 477
                                     :::||||||            :::|||||||||
b       1381 ---------------------GAGAACGATAATGGCGAAGCTGTGATCAAGAAGAGTTTC 1419

y        478 IleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLys 497
             ::::::   ::::::   :::||||||   ||||||      |||   :::   ||||||
b       1420 CTGGAGATAGAGAAGAAGGTTATTACACATTGGAAGGATGTAGTGATGAACGGCAAGAAG 1479

y        498 ThrSerGlyIleCys 502
             |||:::   :::|||
b       1480 ACTAATAAAGTTTGC 1494
```

```
ESULT 11
S-10-006-852-17
 Sequence 17, Application US/10006852
 Publication No. US20030046732A1
 GENERAL INFORMATION:
  APPLICANT: Kinnersely, Alan M.
  APPLICANT:  Turano, Frank J.
  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
  FILE REFERENCE: 7224-65
  CURRENT APPLICATION NUMBER: US/10/006,852
  CURRENT FILING DATE:  2002-07-01
  PRIOR APPLICATION NUMBER: US 60/246,367
  PRIOR FILING DATE: 2000-11-07
  NUMBER OF SEQ ID NOS: 24
  SOFTWARE: PatentIn version 3.1
 SEQ ID NO 17
   LENGTH: 1783
   TYPE: DNA
   ORGANISM: Lycopersicon esculentum
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (6)..(1511)
   OTHER INFORMATION:
S-10-006-852-17
```

```
Alignment Scores:
Pred. No.:                1.46e-224       Length:          1783
Score:                    2014.00         Matches:         387
Percent Similarity:       86.44%          Conservative:    53
Best Local Similarity:    76.03%          Mismatches:      55
Query Match:              77.02%          Indels:          14
DB:                       14              Gaps:            5

US-10-006-852-2 (1-502) x US-10-006-852-17 (1-1783)

Qy          1 MetValLeu---SerHisAlaValSerGluSerAspValSerValHisSerThrPheAla 19
              |||||||||   :::   ::::::   :::|||:::   |||:::|||   |||||||||
Db          6 ATGGTGTTAACAACGACGTCGATAAGAGATTCAGAAGAGAGCTTGCACTGTACATTTGCA 65

Qy         20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
              ||||||||||||:::      ||||||:::||||||||||||:::   |||:::||||||
Db         66 TCAAGATATGTACAGGAACCTTTACCTAAGTTCAAAATGCCTAAAAAATCCATGCCGAAA 125

Qy         40 GluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
              ||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||||
Db        126 GAAGCAGCTTATCAGATTGTAAACGACGAGCTTATGTTGGATGGTAACCCCAGGTTGAAT 185

Qy         60 LeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSer 79
              |||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        186 TTAGCTTCCTTTGTTAGCACATGGATGGAGCCCGAGTGCGATAAGCTCATCATGTCATCC 245

Qy         80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 ATTAATAAAAACTATGTCGACATGGATGAGTATCCTGTCACCACTGAACTTCAAAATAGA 305

Qy        100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaVal 119
              ||||||||||||:::||||||||||||:::||||||:::   :::   ||||||||||||
Db        306 TGTGTTAACATGTTAGCACATCTTTTCCATGCCCCGGTTGGTGATGATGAGACTGCAGTT 365

Qy        120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GGAGTTGGTACAGTGGGTTCATCAGAGGCAATAATGCTTGCTGGCCTTGCTTTCAAACGC 425

Qy        140 LysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThr 159
              |||||||||:::||||||||||||||||||||||||   |||||||||||||||||||||
Db        426 AAATGGCAATCGAAAAGAAAAGCAGAAGGCAAACCTTTCGATAAGCCTAATATAGTCACT 485

Qy        160 GlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLys 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 GGAGCTAATGTGCAGGTCTGCTGGGAAAAATTTGCAAGGTATTTTGAGGTTGAGTTGAAG 545

Qy        180 GluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetVal 199
              ||||||||||||   ||||||||||||||||||||||||   :::||||||::::::|||
Db        546 GAGGTGAAACTAAAAGAAGGATACTATGTAATGGACCCTGCCAAAGCAGTAGAGATAGTG 605

Qy        200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
              ||||||||||||||||||||||||   ||||||||||||||||||   ||||||||||||
Db        606 GATGAGAATACAATATGTGTTGCTGCAATCCTTGGTTCTACTCTGACTGGGGAGTTTGAG 665

Qy        220 AspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThr 239
              ||||||||||||||||||:::||||||   :::|||||||||||||||||||||:::|||
Db        666 GATGTGAAGCTCCTAAACGAGCTCCTTACAAAAAAGAACAAGGAAACCGGATGGGAGACA 725

Qy        240 ProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeu 259
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||:::|||
Db        726 CCGATTCATGTCGATGCTGCGAGTGGAGGATTTATTGCTCCTTTCCTCTGGCCAGATCTT 785

Qy        260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 GAATGGGATTTCCGTTTGCCTCTTGTGAAAAGTATAAATGTCAGCGGTCACAAGTATGGC 845

Qy        280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299
              |||||||||||||||:::||||||||||||||||||:::|||||||||||||||:::|||
Db        846 CTTGTATATGCTGGTGTCGGTTGGGTGATATGGCGGAGCAAGGAAGACTTGCCCGATGAA 905

Qy        300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
              |||:::|||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db        906 CTCGTCTTTCATATAAACTACCTTGGGTCTGATCAGCCTACTTTTACTCTCAACTTCTCT 965
```

```
y         320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
              |||||||||   |||:::|||||||||||||||||||||||||||||||||   ||||||
b         966 AAAGGTTCCTATCAAATAATTGCACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGT 1025

y         340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
              |||:::|||||||||:::||||||      |||      ||||||   ||||||:::
b        1026 TATAAGAACGTCATGAAGAATTGCTTATCAAACGCAAAAGTACTAACAGAGGGAATCACA 1085

y         360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
              |||      ||||||:::|||||||||||||||   |||||||||:::||||||||||||
b        1086 AAAATGGGGCGGTTCGATATTGTCTCTAAGGATGTGGGTGTTCCTGTTGTAGCATTTTCT 1145

y         380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGly 399
              |||:::|||||||||   :::|||   ||||||:::|||:::   |||||||||:::|||
b        1146 CTCAGGGACAGCAGCAAATATACGGTATTTGAAGTATCTGAGCATCTCAGAAGATTTGGA 1205

y         400 TrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgVal 419
              |||||||||||||||||||||||||||:::|||:::|||||||||   ||||||||||||
b        1206 TGGATCGTCCCTGCATACACAATGCCACCGGATGCTGAACACATTGCTGTACTGCGGGTT 1265

y         420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
              |||||||||||||||||||||   :::||||||||||||||||||   ||||||||||||
b        1266 GTCATTAGAGAGGATTTCAGCCACAGCCTAGCTGAGAGACTTGTTTCTGACATTGAGAAA 1325

y         440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIle------SerLeu 457
              ::::::   |||||||||      |||   |||:::      |||
b        1326 ATTCTGTCAGAGTTGGACACACAGCCTCCTCGTTTGCCCACCAAAGCTGTCCGTGTCACT 1385

y         458 GlyGlnGluLysSerGluSerAsnSerAspAsnLeu------------MetValThrVal 473
                 :::|||      :::         |||   |||            |||   ||||||
b        1386 GCTGAGGAAGTGCGTGATGACAAGGGTGATGGGCTTCATCATTTTCACATGGATACTGTA 1445

y         474 LysLysSerAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAla 493
                             :::   |||:::|||||||||      |||:::|||
b        1446 ---------------GAGACTCAGAAAGACATTATCAAACATTGGAGGAAAATCGCAGGG 1490

y         494 AspArgLysLysThrSerGlyIleCys 502
                    |||||||||||||||:::|||
b        1491 ------AAGAAGACCAGCGGAGTCTGC 1511

ESULT 12
S-09-887-576-782
 Sequence 782, Application US/09887576
 Patent No. US20020144047A1
 GENERAL INFORMATION:
  APPLICANT: Budworth, P.
  APPLICANT:  Brown, D.
  APPLICANT:  Chang, H.
  APPLICANT:  Zhu, T.
  APPLICANT:  Han, B.
  APPLICANT:  Wang, X.
  APPLICANT:  Cooper, Bret
  TITLE OF INVENTION: Promoters for regulation of plant expression
  FILE REFERENCE: 1360.001US1
  CURRENT APPLICATION NUMBER: US/09/887,576
  CURRENT FILING DATE:  2001-06-25
  PRIOR APPLICATION NUMBER: US 60/213,848
  PRIOR FILING DATE: 2000-06-23
  PRIOR APPLICATION NUMBER: US 60/214,087
  PRIOR FILING DATE: 2000-06-23
  PRIOR APPLICATION NUMBER: US 60/258,692
  PRIOR FILING DATE: 2000-12-29
  NUMBER OF SEQ ID NOS: 875
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 782
   LENGTH: 1455
   TYPE: DNA
   ORGANISM: Oryza sativa
JS-09-887-576-782

Alignment Scores:
Pred. No.:              6.97e-224      Length:         1455
Score:                      2007.00      Matches:        381
Percent Similarity:         87.67%       Conservative:   60
Best Local Similarity:      75.75%       Mismatches:     42
Query Match:                76.75%       Indels:         20
DB:                         10           Gaps:           6

US-10-006-852-2 (1-502) x US-09-887-576-782 (1-1455)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHis---SerThrPheAla 19
              |||   ||||||   |||      ::::::      |||:::|||   ||||||||||||
Db          1 ATGGCGCTGTCGACGGCG------CAGACAGGGGAGTCGATGCACTCCTCGACGTTCGCG 54

Qy         20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
              ||||||||||||||||||:::||||||||||||:::|||||||||   ||||||||||||
Db         55 TCGCGGTACGTGCGCACGGCGCTGCCGAGGTTCAGGATGCCGGAGAAGTCGATCCCCAAG 114

Qy         40 GluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 GACGCGGCGTACCAGATCATCAACGACGAGCTGATGCTCGACGGCAACCCGCGCCTGAAC 174

Qy         60 LeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSer 79
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||::::::
Db        175 CTGGCGTCCTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATGATGGCCGCC 234

Qy         80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 ATCAACAAGAACTACGTCGACATGGATGAGTACCCCGTCACCACCGAGCTCCAGAACCGG 294

Qy        100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaVal 119
              ||||||||||||||||||||||||||||||||||||:::   :::   ||||||||||||
Db        295 TGCGTGAACATGATCGCGCATCTGTTCAACGCGCCGATCGGGGACGACGAGACGGCGGTC 354

Qy        120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GGGGTGGGCACGGTGGGGTCGTCGGAGGCCATCATGCTGGCGGGGCTGGCGTTCAAGAGG 414

Qy        140 LysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThr 159
              ||||||||||||:::   ||||||||||||||||||   |||||||||||||||||||||
Db        415 AAGTGGCAGAACAGGATGAAGGCCGAGGGGAAGCCCCACGACAAGCCCAACATCGTGACG 474

Qy        160 GlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLys 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GGGGCCAACGTGCAGGTGTGCTGGGAGAAGTTCGCGCGCTACTTCGAGGTGGAGCTCAAG 534

Qy        180 GluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetVal 199
              ||||||||||||::::::|||||||||||||||:::|||::::::||||||:::||||||
Db        535 GAGGTGAAGCTGACCCAAGGGTACTACGTGATGAACCCGGAGAAGGCCGTGGAGATGGTC 594

Qy        200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        595 GACGAGAACACCATCTGCGTCGCCGCCATCCTCGGCTCCACCCTCAACGGCGAGTTCGAG 654

Qy        220 AspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThr 239
              |||||||||:::|||||||||||||||      ||||||   ||||||||||||:::|||
Db        655 GACGTCAAGATGCTCAACGACCTCCTCACCGCCAAGAACGCCGAGACAGGGTGGAACACG 714

Qy        240 ProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeu 259
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        715 CCGATCCATGTGGACGCGGCGAGCGGCGGGTTCATCGCGCCGTTCATCTACCCGGAGCTG 774

Qy        260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 GAGTGGGACTTCCGGCTGCCGCTGGTGAAGAGCATCAACGTCAGCGGCCACAAGTACGGG 834

Qy        280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299
              |||||||||||||||:::||||||||||||||||||||||||||||||||||||:::|||
Db        835 CTCGTCTACGCCGGCGTCGGGTGGGTCATCTGGCGCAACAAGGAGGACCTCCCCGATGAG 894

Qy        300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        895 CTCATCTTCCACATCAACTACCTCGGCGCCGACCAGCCAACCTTCACGCTCAACTTCTCC 954

Qy        320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
```

```
            |||||||||:::|||:::|||||||||||||||||||||||||||||||||   ||||||
)        955 AAAGGATCGAACCAGATAATTGCGCAGTATTACCAGCTCATTCGTCTCGGATTCGAGGGG 1014

'        340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
            |||:::::::::|||:::|||||||||:::|||      |||||||||||||||:::|||
)       1015 TACAAGGACATCATGCAGAACTGCCGGGACAACGCGACGGTGCTCCGGGAGGGGATCGAG 1074

'        360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
            ||||||      |||::::::||||||||||||   ||||||||||||||||||||||||
)       1075 AAGACGGGCCACTTCGACGTGGTGTCCAAGGACTCCGGCGTGCCGCTGGTGGCCTTCTCC 1134

'        380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGly 399
            |||||||||||||||   :::|||   ||||||:::::::::   |||||||||:::|||
        1135 CTCAAGGACTCGTCGCGGTACACGGTGTTCGAGGTGGCCGAGAGCCTCCGCCGCTTCGGC 1194

'        400 TrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgVal 419
            |||||||||||||||||||||||||||   :::|||:::|||:::   |||:::||||||
)       1195 TGGATCGTGCCGGCGTACACCATGCCCGCCGACGCTGAGCACGTCGCCGTGATGCGCGTC 1254

'        420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
            ||||||||||||||||||||||||   |||||||||||||||:::   |||:::   |||
)       1255 GTCATCCGCGAGGACTTCAGCCGCGGCCTCGCCGAGCGCCTCATCACCGACCTCACCAAG 1314

'        440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGln 459
               :::   ::::::|||                     |||   :::
)       1315 ACGGTGGCCGATATGGAC------------------GCCCACGCCGTC------------ 1344

y        460 GluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAsp 479
            :::|||:::   :::                           ||||||:::      :::
)       1345 AAGAAGGCCGCCGCCGAG------------------CCGGCCAAGAAGACCGTGCGGGAG 1386

y        480 LysGlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSer 499
               ::::::::::::   |||   |||:::   |||||||||   |||||||||:::|||
)       1387 ATAGAGAAGGAGGTGACCACCTACTGGCGGAGTTTCGTCGCC---AGGAAGAAGAGCAGC 1443

y        500 GlyIleCys 502
               :::|||
b       1444 CTCGTCTGC 1452

ESULT 13
S-10-006-852-7
 Sequence 7, Application US/10006852
 Publication No. US20030046732A1
 GENERAL INFORMATION:
  APPLICANT: Kinnersely, Alan M.
  APPLICANT:  Turano, Frank J.
  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
  FILE REFERENCE: 7224-65
  CURRENT APPLICATION NUMBER: US/10/006,852
  CURRENT FILING DATE:  2002-07-01
  PRIOR APPLICATION NUMBER: US 60/246,367
  PRIOR FILING DATE: 2000-11-07
  NUMBER OF SEQ ID NOS: 24
  SOFTWARE: PatentIn version 3.1
 SEQ ID NO 7
   LENGTH: 2121
   TYPE: DNA
   ORGANISM: Arabidopsis thaliana
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (1)..(87)
   OTHER INFORMATION:
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (274)..(681)
   OTHER INFORMATION:
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (782)..(994)
   OTHER INFORMATION:
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (1081)..(1335)
```

```
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1438)..(1488)
;   OTHER INFORMATION:
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1654)..(2118)
;   OTHER INFORMATION:
US-10-006-852-7

Alignment Scores:
Pred. No.:              2.13e-215      Length:         2121
Score:                  1936.50        Matches:        410
Percent Similarity:     61.99%         Conservative:   32
Best Local Similarity:  57.50%         Mismatches:     46
Query Match:            74.05%         Indels:         226
DB:                     14             Gaps:           7

US-10-006-852-2 (1-502) x US-10-006-852-7 (1-2121)

Qy           1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
               ||||||||||||      |||||||||||||||||||||:::||||||||||||||||||
Db           1 ATGGTTTTGTCTAAGACAGTTTCCGAATCTGATGTCTCAATCCATTCAACTTTTGCTTCT 60

Qy          21 ArgTyrValArgThrSerLeuPro------------------------------------ 28
               ||||||||||||   |||||||||
Db          61 CGTTACGTCCGCAACTCTCTTCCACGGTAACAACTTGTAACACAAATCTTTTGTCTAATG 120

Qy          28 ------------------------------------------------------------ 28

Db         121 TTTTCGTCAACAATAGTAACATGTAATGATGTAAACCTTGGATAGTTTTTTTTTGGCCG 180

Qy          28 ------------------------------------------------------------ 28

Db         181 TGGTTAATGTTGTAGATTTATTATGTGTTATATACTATAAGGAAGGACATGTTTCGTTAT 240

Qy          29 ------------------------------ArgPheLysMetProGluAsnSerIlePro 38
                                             ||||||:::|||||||||||||||||||||
Db         241 TTTAACTTAATGTATCATCATTTCATCATTAGATTCGAAATGCCTGAGAACTCAATCCCA 300

Qy          39 LysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeu 58
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 AAAGAAGCAGCTTACCAAATCATCAACGACGAGCTAATGCTCGATGGTAACCCAAGGCTG 360

Qy          59 AsnLeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSer 78
               |||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db         361 AACCTAGCTTCCTTCGTGACCACATGGATGGAGCCAGAATGTGACAAGCTCATGATGGAG 420

Qy          79 SerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsn 98
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TCCATCAACAAGAACTACGTCGACATGGACGAGTACCCTGTCACCACTGAGCTTCAGAAC 480

Qy          99 ArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAla 118
               |||||||||||||||||||||   ||||||||||||||||||   :::   |||   |||
Db         481 CGATGTGTTAACATGATAGCACGTCTCTTCAACGCGCCGCTTGGTGACGGTGAAGCTGCC 540

Qy         119 ValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLys 138
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GTTGGTGTTGGCACCGTCGGATCGTCGGAGGCGATTATGTTGGCCGGTTTGGCTTTTAAG 600

Qy         139 ArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleVal 158
               |||:::|||||||||||||||||||||:::|||   |||   ||||||||||||||||||
Db         601 AGACAATGGCAGAATAAGCGTAAGGCCCAAGGGCTTCCTTATGATAAGCCCAATATCGTA 660

Qy         159 ThrGlyAlaAsnVal--------------------------------------------- 163
               |||||||||||||||
Db         661 ACCGGTGCTAATGTCCAGGTAAACCAAAACAAAAATTGATGAAATATTAACCAAGACAAA 720

Qy         164 -------------------------------------------------------GlnVa 165
                                                                      |||||
Db         721 ATTGAATTTATCAATCCGGTTAAGTTATATGTGTGACTCAATTTCCGGTTCAATACAGGT 780
```

```
y        165 lCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSerGl 185
             |||||||||||||||||||||||||||||||||||||||||||||||||   |||   ||
b        781 TTGCTGGGAGAAATTCGCAAGGTATTTCGAAGTGGAGCTTAAGGAAGTGAACCTAAGAGA 840

y        185 uGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIleCy 205
             |   ||||||||||||||||||   :::||||||:::|||||||||||||||||||||||
b        841 AGACTATTACGTGATGGACCCTGTAAAGGCGGTCGAAATGGTAGACGAAAACACAATTTG 900

y        205 sValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspValLysLeuLeuAs 225
             |||||||   ||||||||||||||||||   |||||||||||||||||||||||||||||
b        901 TGTCGCTGCCATCCTCGGTTCAACGTTAACCGGTGAATTCGAAGACGTTAAGCTCCTCAA 960

y        225 nAspLeuLeuValGluLysAsnLysGluThrGly------------------------- 236
             |||||||||||||||||||||||||:::|||
b        961 CGACCTCCTTGTCGAGAAAAACAAGCAAACCGG-GTAATTAAACCAAACCGAGAAACAAG 1019

y        236 ------------------------------------------------------------ 236

b       1020 CTAATATCGATTGTAATCGGTTTGGAGTCCGGTTTTAACGTTCTAAAACACAATTTGCAG 1079

y        237 -TrpAspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTy 256
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       1080 ATGGGACACGCCAATACACGTGGACGCAGCGAGTGGTGGGTTTATTGCTCCGTTCTTGTA 1139

y        256 rProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHi 276
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       1140 TCCGGAGCTGGAGTGGGATTTCCGGCTACCGTTGGTTAAGAGTATTAATGTGAGTGGTCA 1199

y        276 sLysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLe 296
             |||||||||||||||||||||||||||||||||||||:::||||||   |||   |||||
b       1200 CAAATACGGTTTGGTTTACGCCGGTATTGGTTGGGTTGTATGGAGAACCAAAACCGATTT 1259

y        296 uProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLe 316
             ||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||
b       1260 GCCTGATGAACTTATCTTCCATATCAATTATCTTGGCGCTGATCAACCAACCTTTACACT 1319

y        316 uAsnPheSerLys----------------------------------------------- 320
             |||||||||||||
b       1320 CAACTTCTCCAAAGGTACATTACCATAAGTCCATAACATATATAACTTTCAATAATATTT 1379

y        321 -------------------------------------------------------GlySe 322
                                                                    |||||
b       1380 TTGGTGTATGGAATTGTTTTATAGACTAAACATTTGATAATGCTTGTATAAACCAGGTTC 1439

y        322 rSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGlu----------- 338
             |||||||||||||||||||||||||||||||||||||||||||   |||
b       1440 AAGTCAAGTGATTGCTCAGTACTACCAGCTGATTCGTCTTGGATTCGAGGTAAATAATAA 1499

y        338 ------------------------------------------------------------ 338

b       1500 CTCAAAATAGCAATATATTTACCAAATGGTCAATAAAGAAACTAGAATGTATTATATTTA 1559

y        338 ------------------------------------------------------------ 338

b       1560 AGTTGTTACTTGTTACTATACTTTGAATTAAACGTTCCTAACATGACTAGTTTTGGTATT 1619

y        339 ----------------------------------GlyTyrArgAsnValMetGluAsnCy 347
                                               ||||||||||||||||||:::|||||
b       1620 GTGTAATTAATAATGTTTTTCTTGTTTGATTTAGGGTTATCGCAATGTGATGGATAATTG 1679

y        347 sArgGluAsnMetIleValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleVa 367
             |||||||||||||:::|||||||||:::|||||||||||||||   ||||||   |||||
b       1680 TCGGGAAAACATGATGGTACTAAGACAAGGATTAGAGAAAACGGGACGTTTTAAAATCGT 1739

y        367 lSerLysAspGluGlyValProLeuValAlaPheSerLeuLysAspSerSerCysHisTh 387
             |||||||:::   ||||||||||||||||||||||||||||||||||||||    |||
b       1740 CTCCAAAGAAAACGGTGTTCCGTTAGTGGCGTTTTCTCTCAAAGATAGTAGCCGCCACAA 1799

y        387 rGluPheGluIleSerAspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMe 407
              |||||||||::::::      |||||||||:::||||||||||||||||||||||||||
b       1800 CGAGTTCGAGGTGGCCCATACACTCCGTCGCTTCGGCTGGATCGTTCCGGCCTACACGAT 1859

y        407 tProProAsnAlaGlnHisIleThrValLeuArgValValIleArgGluAspPheSerAr 427
             ||||   :::|||||||||:::||||||||||||||||||||||||||||||||||||||
Db      1860 GCCTGCGGATGCGCAGCATGTCACTGTCCTTCGAGTTGTTATCCGAGAAGATTTCTCTCG 1919

Qy       427 gThrLeuAlaGluArgLeuValIleAspIleGluLysValMetArgGluLeuAspGluLe 447
             ||||||||||||||||||||||   |||   |||||||||:::   |||||||||   ||
Db      1920 AACCTTAGCCGAGAGATTGGTAGCTGATTTCGAGAAGGTTCTACACGAGCTCGATACGCT 1979

Qy       447 uProSerArgValIleHisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSerAs 467
             ||||:::||||||      |||::::::   |||:::
Db      1980 TCCGGCGAGGGTTCACGCCAAGATGGCTAATGGAAAAGTTAACGGT-------------- 2025

Qy       467 pAsnLeuMetValThrValLysLysSerAspIleAspLysGlnArgAspIleIleThrGl 487
                             |||||||||:::      :::   ||||||::::::
Db      2026 ----------------GTTAAGAAGACGCCAGAGGAGACGCAGAGAGAAGTCACGGCCTA 2069

Qy       487 yTrpLysLysPheValAlaAspArgLysLysThrSer 499
              |||||||||   :::   :::   |||||||||:::
Db      2070 CTGGAAGAAGTTGTTG---GAGACTAAGAAGACCAAC 2103

RESULT 14
US-09-887-576-806
; Sequence 806, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
;  APPLICANT: Budworth, P.
;  APPLICANT:  Brown, D.
;  APPLICANT:  Chang, H.
;  APPLICANT:  Zhu, T.
;  APPLICANT:  Han, B.
;  APPLICANT:  Wang, X.
;  APPLICANT:  Cooper, Bret
;  TITLE OF INVENTION: Promoters for regulation of plant expression
;  FILE REFERENCE: 1360.001US1
;  CURRENT APPLICATION NUMBER: US/09/887,576
;  CURRENT FILING DATE:  2001-06-25
;  PRIOR APPLICATION NUMBER: US 60/213,848
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/214,087
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/258,692
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 875
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 806
;   LENGTH: 1287
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(1287)
;   OTHER INFORMATION: n = A,T,C or G
US-09-887-576-806

Alignment Scores:
Pred. No.:              2.97e-210      Length:         1287
Score:                  1889.50        Matches:        362
Percent Similarity:     91.47%         Conservative:   24
Best Local Similarity:  85.78%         Mismatches:     20
Query Match:            72.26%         Indels:         17
DB:                     10             Gaps:           1

US-10-006-852-2 (1-502) x US-09-887-576-806 (1-1287)

Qy        32 MetProGluAsnSerIleProLysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMet 51
             |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ATGCCGGAGCAGTCGATCCCCAAGGAGGCGGCGTACCAGATCATCAACGACGAGCTGATG 60

Qy        52 LeuAspGlyAsnProArgLeuAsnLeuAlaSerPheValThrThrTrpMetGluProGlu 71
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CTGGACGGCAACCCGCGGCTGAACCTCGCGTCGTTCGTCACCACGTGGATGGAGCCCGAG 120

Qy        72 CysAspLysLeuIleMetSerSerIleAsnLysAsnTyrValAspMetAspGluTyrPro 91
             |||||||||||||||   :::|||:::|||||||||||||||||||||||||||||||||
```

```
b          121 TGCGACAAGCTCATCCAGGCCTCCGTCAACAAGAACTACGTCGACATGGACGAGTACCCC 180

y           92 ValThrThrGluLeuGlnAsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaPro 111
               ||||||||||||||||||
b          181 GTCACCACCGAACTCCAGGG---------------------------------------- 200

y          112 LeuGluGluAlaGluThrAlaValGly-ValGlyThrValGlySerSerGluAlaIleMe 131
                     ::::::||||||||||||||| ||||||||||||||||||||||||||||||||
b          201 ------GACTCTGAAACGGCCGTCGGAAGTCGGCACTGTCGGCTCGTCTGAGGCCATCAT 254

y          131 tLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLysArgLysAlaGluGlyLysPr 151
               |||||||||||||||||||||||||:::||||||||||||   ||||||   ||||||||
b          255 GCTCGCCGGTTTGGCCTTCAAGAGGAGGTGGCAGAACAAGATGAAGGCAGCCGGCAAGCC 314

y          151 oValAspLysProAsnIleValThrGlyAlaAsnValGlnValCysTrpGluLysPheAl 171
               |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          315 ATGCGACAAGCCTAACATTGTCACCGGCGCCAATGTCCAAGTTTGCTGGGAGAAGTTCGC 374

y          171 aArgTyrPheGluValGluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAs 191
               ||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||
b          375 GCGATACTTCGAGGTTGAGCTCAAGGAAGTGAAGCTGAGTGACGGCTACTACGTCATGGA 434

y          191 pProGlnGlnAlaValAspMetValAspGluAsnThrIleCysValAlaAspIleLeuGl 211
               ||||   :::|||||||||||||||||||||||||||||||||||||||   ||||||||
b          435 CCCAGCTAAGGCCGTGGATATGGTCGACGAGAACACCATCTGCGTCGCGGCGATCCTCGG 494

y          211 ySerThrLeuAsnGlyGluPheGluAspValLysLeuLeuAsnAspLeuLeuValGluLy 231
               ||||||||||||||||||||||||||||||||||||||||||||||||||||   :::||
b          495 GTCGACGCTGAACGGGGAGTTCGAGGACGTGAAGCTGCTCAACGATCTGCTCACCAAGAA 554

y          231 sAsnLysGluThrGlyTrpAspThrProIleHisValAspAlaAlaSerGlyGlyPheIl 251
               ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
b          555 GAACGCTGAAACAGGCTGGGACACGCCGATCCACGTGGACGCGGCGAGCGGCGGGTTCAT 614

y          251 eAlaProPheLeuTyrProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIl 271
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          615 CGCGCCGTTCCTGTACCCGGAGCTGGAGTGGGACTTCCGGCTGCCGCTGGTGAAGAGCAT 674

y          271 eAsnValSerGlyHisLysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpAr 291
               |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||
b          675 CAACGTGAGCGGGCACAAGTACGGCCTCGTCTACGCCGGGATCGGGTGGTGCATCTGGAG 734

y          291 gAsnLysGluAspLeuProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGl 311
               |:::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          735 GAGCAAGGAGGATCTGCCTGAGGAGCTCATCTTCCACATCAACTACCTCGGCGCCGACCA 794

Qy         311 nProThrPheThrLeuAsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGl 331
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         795 GCCCACCTTCACCCTCAACTTCTCCAAGGGTTCCAGCCAGGTCATTGCACAGTATTACCA 854

Qy         331 nLeuIleArgLeuGlyHisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMe 351
               ||||||||||||||||   |||||||||:::|||:::||||||||||||:::||||||
Db         855 ACTAATCCGCCTAGGCTTTGAGGGGTACAAGAACATCATGGAGAACTGCCAGGAGAACGC 914

Qy         351 tIleValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGl 371
                :::||||||::::::|||||||||||||||   ||||||||||||||||||||||||
Db         915 GATGGTGCTGAAGCAGGGGCTGGAGAAGACGGGGCGGTTCAACATCGTGTCCAAGGACAA 974

Qy         371 uGlyValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIl 391
                ||||||||||||||||||||||||||||||||||||:::   |||   |||||||||||
Db         975 CGGCGTGCCGCTGGTGGCCTTCTCCCTCAAGGACAGCGCCCGGCACAACGAGTTCGAGAT 1034

Qy         391 eSerAspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAl 411
               |||||||   |||||||||:::|||||||||||||||||||||||||||||||||:::||
Db        1035 CTCCGACTTCCTCCGCCGCTTCGGCTGGATCGTGCCGGCCTACACCATGCCCCCCGACGC 1094

Qy         411 aGlnHisIleThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAlaGl 431
               |||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1095 GCAGCACGTCACCGTGCTCCGCGTCGTCATCCGCGAGGACTTCAGCCGCACGCTCGCCGA 1154

Qy         431 uArgLeuValIleAspIleGluLysValMetArgGluLeuAspGluLeuProSerArgVa 451
               ||||||||||:::|||:::|||||||||:::   |||||||||   ||||||:::|||||
Db        1155 GCGCCTCGTGCTCGACGTCGAGAAGGTGCTGCACGAGCTCGACGCGCTCCCSGCCCGCGT 1214
```

```
Qy         451 lIle 452
               |:::
Db        1215 CGTC 1218

RESULT 15
US-10-167-547C-3
; Sequence 3, Application US/10167547C
; Publication No. US20030170653A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  APPLICANT:  Damude, Howard G.
;  TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
;  TITLE OF INVENTION:  Butyrolactone and its Intermediates
;  FILE REFERENCE: CL1804 US NA
;  CURRENT APPLICATION NUMBER: US/10/167,547C
;  CURRENT FILING DATE:  2003-03-17
;  PRIOR APPLICATION NUMBER: 60/297198
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: Microsoft Office 07
; SEQ ID NO 3
;   LENGTH: 1530
;   TYPE: DNA
;   ORGANISM: Alstroemeria
US-10-167-547C-3

Alignment Scores:
Pred. No.:                3.69e-207      Length:         1530
Score:                    1864.00        Matches:        351
Percent Similarity:       84.55%         Conservative:   54
Best Local Similarity:    73.28%         Mismatches:     70
Query Match:              71.28%         Indels:         4
DB:                       12             Gaps:           2

US-10-006-852-2 (1-502) x US-10-167-547C-3 (1-1530)

Qy           1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
               |||   ||||||      ||||||:::|||:::      |||      |||:::||||||
Db           1 ATGGCTCTCTCCAGCGTCGTCTCCGACTCCAACAACCAAGTGCAGTGCACCTATGCCTCT 60

Qy          21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
               ||||||||||||         |||   |||:::|||||||||   |||||||||||||||
Db          61 CGCTACGTTCGCGACGAGGCTCCAGGGTTCAGGATGCCGGAGAAGTCGATACCAAAGGAG 120

Qy          41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
               ||||||:::   :::|||||||||||||||||||||||||||||||||||||||||||||
Db         121 GCGGCGTTCACGATGATCAACGACGAGCTGATGCTGGACGGGAACCCCAGGCTGAACTTG 180

Qy          61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
               |||||||||||||||||||||||||||||||||||||||:::|||:::||||||:::|||
Db         181 GCTTCGTTCGTGACGACGTGGATGGAGCCGGAGTGCGATCGTCTGATGATGTCCACCATC 240

Qy          81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
               ||||||||||||      ||||||:::||||||||||||   ::::::||||||||||||
Db         241 AACAAGAACTACGCCCTCATGGACGATTACCCGGTCACTATTGACATACAGAATCGCTGC 300

Qy         101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
               |||||||||||||||:::|||||||||||||||:::   |||   ||||||   ||||||
Db         301 GTGAATATGATAGCCAACCTCTTTAATGCGCCAATTGGGGAGGGGGAAACAACAGTAGGA 360

Qy         121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
                     |||||||||||||||||||||:::|||||||||||||||||||||||||||
Db         361 TGTGCTACGGTGGGATCATCAGAAGCCATGATGCTTGCAGGGTTGGCATTCAAGAGAAAT 420

Qy         141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
               |||||||||||||||||||||||||||||||||   ||||||||||||:::|||||||||
Db         421 TGGCAGAACAAAAGAAAGGCAGAGGGGAAGCCATATGACAAGCCCAACATGGTCACCGGT 480

Qy         161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
               :::||||||||||||||||||   |||||||||:::|||||||||||||||:::|||:::
Db         481 TCAAATGTTCAGGTTTGCTGGGTGAAATTCGCTAAGTATTTTGAAGTTGAAATGAAAAAA 540
```

```
y       181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
            |||   |||   ||||||||||||||||||||||||::::::||||||:::|||||||||
b       541 GTGAATTTGAGGGAGGGATATTATGTGATGGACCCAGAGAAGGCTGTGGAAATGGTGGAT 600

y       201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
            |||||||||||||||||||||   |||||||||||||||||||||   ||||||||||||
b       601 GAGAATACCATTTGTGTTGCTGCCATCTTGGGCTCGACCCTTACTGGAGAGTTCGAAGAT 660

y       221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
            |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
b       661 GTCAAACTGCTAAACGACCTCCTCGTAGAGAAGAACAAGAAAACTGGTTGGGATACACCC 720

y       241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
            ||||||||||||||||||   |||||||||||||||||||||:::|||||||||||||||
b       721 ATTCATGTTGATGCTGCTATTGGTGGATTCATTGCTCCATTTATCTATCCAGAACTGGAA 780

y       261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       781 TGGGATTTCCGACTACCTCTGGTGAAGAGTATCAATGTCAGTGGACACAAATACGGCCTT 840

y       281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
            ||||||   |||:::|||||||||:::||||||||||||   ||||||||||||||||||
b       841 GTCTATCCCGGCGTTGGTTGGGTCGTCTGGAGGAACAAGAATGATCTTCCTGAAGAACTC 900

y       301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
            ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
b       901 ATTTTCCATATCAACTATCTTGGGATTGATCAACCCACTTTTACCCTCAACTTCTCGAAA 960

y       321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
            ||||||:::|||:::|||   |||||||||||||||||||||||||||   |||||||||
b       961 GGTTCAAACCAGATAATTGGTCAATACTATCAATTAATTCGCCTTGGTTTTGAGGGGTAC 1020

y       341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
            :::|||:::||||||||||||   ||||||      :::|||||||||   ||||||:::
b      1021 AAGAATATAATGGAAAACTGCACGGAGAATGCAAGAATTCTTAGGGAACATCTCGAGGAG 1080

y       361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
                     |||   |||:::|||||||||   |||   |||||||||      :::|||
b      1081 ATGGGCGTTTTCGAGATCATCTCCAAGGATATTGGGGCGCCTCTCGTCACAATTGCTCTC 1140

y       381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
            ||||||||||||   |||:::   |||:::|||:::|||   :::||||||:::||||||
b      1141 AAGGACAGCAGCAAACATAGTGTCTTTAAGATAGCCGATACAATTAGAAGGTTTGGATGG 1200

y       401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
               :::||||||||||||||||||   :::   :::||||||   |||||||||||||||
b      1201 ACAATTCCTGCATACACAATGCCAAAGGACGTTGAGCACATAGCCGTCCTTCGTGTGGTT 1260

y       421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
            |||||||||||||||||||||:::|||||||||||||||      |||::::::||||||
b      1261 ATCAGGGAGGACTTCAGCCGGAGCCTCGCCGAGCGCCTAGCTAATGACATGAAGAAGGTG 1320

y       441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
            :::   |||||||||      |||||||||               ::::::
b      1321 TTGGTTGAGCTGGACATACATCCCAGTCGCACC---------ACCACCATTGCCCACGTT 1371

y       461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAsp 479
            |||:::      |||   :::            |||   :::|||||||||   ::::::
b      1372 AAGGCAGTGGAGAATGGCAATGGTAATTATGTG---ATCAAGAAGAGTATTGTAGAG 1425

Search completed: October 22, 2003, 18:41:35
Job time : 404 secs
```

```
                       GenCore version 5.1.6
              Copyright (c) 1993 - 2003  Compugen Ltd.

1 protein - nucleic search, using frame_plus_p2n model

in on:          October 22, 2003, 15:16:40 ; Search time 97 Seconds
                                            (without alignments)
                                            2284.271 Million cell updates/sec

itle:           US-10-006-852-2
erfect score:   2615
equence:        1 MVLSHAVSESDVSVHSTFAS..........DIITGWKKFVADRKKTSGIC 502

coring table:   BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

earched:        569978 seqs, 220691566 residues

otal number of hits satisfying chosen parameters:      1139956

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

ommand line parameters:
MODEL=frame+_p2n.model -DEV=xlp
Q=/cgn2_1/USPTO_spool_p/US10006852/runat_21102003_164104_2478/app_query.fasta_1.647
DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US10006852_@CGN_1_1_103_@runat_21102003_164104_2478 -NCPU=6 -ICPU=3
NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :       Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                %
esult           Query
  No.   Score   Match Length DB  ID                          Description
-----------------------------------------------------------------------------
    1    2014    77.0   1783  2  US-08-522-421-8             Sequence 8, Appli
c   2    1114    42.6 4403765  3  US-09-103-840A-2            Sequence 2, Appli
c   3    1114    42.6 4411529  3  US-09-103-840A-1            Sequence 1, Appli
    4    1025    39.2   5565  3  US-09-068-195-21            Sequence 21, Appl
c   5     605    23.1  36941  4  US-08-311-731A-130          Sequence 130, App
    6     325    12.4   1770  4  US-08-939-309-7             Sequence 7, Appli
    7     325    12.4   1770  4  US-09-849-180-7             Sequence 7, Appli
    8     325    12.4   1770  4  US-09-356-643B-1            Sequence 1, Appli
    9   265.5    10.2   1629  4  US-08-939-309-5             Sequence 5, Appli
   10   265.5    10.2   1629  4  US-09-849-180-5             Sequence 5, Appli
   11   265.5    10.2   1629  4  US-09-356-643B-3            Sequence 3, Appli
c  12     257     9.8    359  4  US-09-056-556-192           Sequence 192, App
```

```
c  13     257     9.8    359  4  US-09-072-596-187           Sequence 187, App
   14   246.5     9.4   1707  4  US-08-939-309-1             Sequence 1, Appli
   15   246.5     9.4   1707  4  US-09-849-180-1             Sequence 1, Appli
   16   246.5     9.4   1707  4  US-09-356-643B-5            Sequence 5, Appli
   17     227     8.7   1707  4  US-08-939-309-3             Sequence 3, Appli
   18     227     8.7   1707  4  US-09-849-180-3             Sequence 3, Appli
   19     227     8.7   1707  4  US-09-356-643B-7            Sequence 7, Appli
   20     225     8.6 1664976  4  US-08-916-421B-1            Sequence 1, Appli
   21     224     8.6   2130  4  US-09-740-369-1             Sequence 1, Appli
   22     177     6.8   3162  4  US-09-356-643B-12           Sequence 12, Appl
   23   160.5     6.1   1467  4  US-08-939-309-9             Sequence 9, Appli
   24   160.5     6.1   1467  4  US-09-849-180-9             Sequence 9, Appli
   25   160.5     6.1   1467  4  US-09-356-643B-9            Sequence 9, Appli
   26   133.5     5.1   1578  4  US-09-328-352-366           Sequence 366, App
c  27   121.5     4.6 1830121  4  US-09-557-884-1             Sequence 1, Appli
c  28   121.5     4.6 1830121  4  US-09-643-990A-1            Sequence 1, Appli
   29     116     4.4   1944  3  US-09-377-557-19            Sequence 19, Appl
   30     113     4.3   1197  4  US-09-328-352-381           Sequence 381, App
c  31   110.5     4.2 129908  4  US-09-585-858-1             Sequence 1, Appli
   32     105     4.0   2190  4  US-09-134-001C-1891         Sequence 1891, Ap
c  33     105     4.0   5340  4  US-09-627-122-21            Sequence 21, Appl
c  34     105     4.0 1664976  4  US-08-916-421B-1            Sequence 1, Appli
   35     104     4.0   3312  4  US-09-669-751-259           Sequence 259, App
   36     104     4.0   4032  1  US-08-107-748-3             Sequence 3, Appli
   37     104     4.0   4032  1  US-08-245-809-4             Sequence 4, Appli
   38     104     4.0   4032  5  PCT-US92-01385-3            Sequence 3, Appli
   39   102.5     3.9   2193  3  US-09-398-865A-1            Sequence 1, Appli
   40   102.5     3.9   2193  4  US-09-710-714-1             Sequence 1, Appli
c  41   102.5     3.9   2992  1  US-07-718-575-1             Sequence 1, Appli
c  42   102.5     3.9   2992  1  US-08-481-206-1             Sequence 1, Appli
c  43   102.5     3.9   2992  2  US-08-486-269A-1            Sequence 1, Appli
   44     102     3.9   1734  3  US-08-687-590-60            Sequence 60, Appl
   45   101.5     3.9   3666  4  US-09-107-532A-305          Sequence 305, App

                                  ALIGNMENTS

RESULT 1
US-08-522-421-8
; Sequence 8, Application US/08522421
; Patent No. 5908973
;  GENERAL INFORMATION:
;    APPLICANT:  Abu-Bakar, Umi Kalsom
;    APPLICANT:  Barton, Sarah Louise
;    APPLICANT:  Gallego-Veigas, Pedro Pablo
;    APPLICANT:  Gray, Julie Elizabeth
;    APPLICANT:  Grierson, Donald
;    APPLICANT:  Lowe, Alexandra Louise
;    APPLICANT:  Picton, Steve
;    APPLICANT:  Whotton, Lee Colin
;    TITLE OF INVENTION:  DNA, DNA CONSTRUCTS, CELLS AND PLANTS
;    TITLE OF INVENTION:  DERIVED THEREFROM
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pillsbury Madison & Sutro, L.L.P.
;      STREET:  1100 New York Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3918
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Microsoft Word
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/522,421
;      FILING DATE:  11-JAN-1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9305868.3
;      FILING DATE:  22-MAR-1993
;    PRIOR APPLICATION DATA:
```

```
    APPLICATION NUMBER:  GB 9305869.1
    FILING DATE:  22-MAR-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  GB 9305859.2
    FILING DATE:  22-MAR-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  GB 9305865.9
    FILING DATE:  22-MAR-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  GB 9305866.7
    FILING DATE:  22-MAR-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  GB 9305867.5
    FILING DATE:  22-MAR-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  GB 9305860.0
    FILING DATE:  22-MAR-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  GB 9305862.6
    FILING DATE:  22-MAR-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  GB 9314351.9
    FILING DATE:  12-JUL-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  GB 9320988.0
    FILING DATE:  12-OCT-1993
 INFORMATION FOR SEQ ID NO:  8:
  SEQUENCE CHARACTERISTICS:
    LENGTH:  1783 base pairs
    TYPE:  nucleic acid
    STRANDEDNESS:  single
    TOPOLOGY:  linear
  MOLECULE TYPE:  cDNA
  ORIGINAL SOURCE:
    ORGANISM:  ERTD1
S-08-522-421-8

lignment Scores:
red. No.:               3.64e-239     Length:         1783
core:                   2014.00       Matches:        387
ercent Similarity:      86.44%        Conservative:   53
est Local Similarity:   76.03%        Mismatches:     55
uery Match:             77.02%        Indels:         14
B:                      2             Gaps:           5

S-10-006-852-2 (1-502) x US-08-522-421-8 (1-1783)

y            1 MetValLeu---SerHisAlaValSerGluSerAspValSerValHisSerThrPheAla 19
               |||||||||   :::   ::::::   :::|||:::   |||:::|||   |||||||||
b            6 ATGGTGTTAACAACGACGTCGATAAGAGATTCAGAAGAGAGCTTGCACTGTACATTTGCA 65

y           20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
               ||||||||||||:::      ||||||:::||||||||||||:::   |||:::||||||
b           66 TCAAGATATGTACAGGAACCTTTACCTAAGTTCAAAATGCCTAAAAAATCCATGCCGAAA 125

y           40 GluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
               ||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||||
b          126 GAAGCAGCTTATCAGATTGTAAACGACGAGCTTATGTTGGATGGTAACCCCAGGTTGAAT 185

y           60 LeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSer 79
               |||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
b          186 TTAGCTTCCTTTGTTAGCACATGGATGGAGCCCGAGTGCGATAAGCTCATCATGTCATCC 245

y           80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          246 ATTAATAAAAACTATGTCGACATGGATGAGTATCCTGTCACCACTGAACTTCAAAATAGA 305

y          100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaVal 119
               ||||||||||||:::||||||||||||:::||||||:::   :::   ||||||||||||
b          306 TGTGTTAACATGTTAGCACATCTTTTCCATGCCCCGGTTGGTGATGATGAGACTGCAGTT 365

y          120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         366 GGAGTTGGTACAGTGGGTTCATCAGAGGCAATAATGCTTGCTGGCCTTGCTTTCAAACGC 425

Qy         140 LysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThr 159
               |||||||||:::||||||||||||||||||||||||   |||||||||||||||||||||
Db         426 AAATGGCAATCGAAAAGAAAAGCAGAAGGCAAACCTTTCGATAAGCCTAATATAGTCACT 485

Qy         160 GlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLys 179
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         486 GGAGCTAATGTGCAGGTCTGCTGGGAAAAATTTGCAAGGTATTTTGAGGTTGAGTTGAAG 545

Qy         180 GluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetVal 199
               ||||||||||||   ||||||||||||||||||||||||   :::||||||::::::|||
Db         546 GAGGTGAAACTAAAAGAAGGATACTATGTAATGGACCCTGCCAAAGCAGTAGAGATAGTG 605

Qy         200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
               ||||||||||||||||||||||||   ||||||||||||||||||   ||||||||||||
Db         606 GATGAGAATACAATATGTGTTGCTGCAATCCTTGGTTCTACTCTGACTGGGGAGTTTGAG 665

Qy         220 AspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThr 239
               ||||||||||||||||||:::||||||   :::|||||||||||||||||||||:::|||
Db         666 GATGTGAAGCTCCTAAACGAGCTCCTTACAAAAAAGAACAAGGAAACCGGATGGGAGACA 725

Qy         240 ProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeu 259
               ||||||||||||||||||||||||||||||||||||||||||||||||:::|||:::|||
Db         726 CCGATTCATGTCGATGCTGCGAGTGGAGGATTTATTGCTCCTTTCCTCTGGCCAGATCTT 785

Qy         260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         786 GAATGGGATTTCCGTTTGCCTCTTGTGAAAAGTATAAATGTCAGCGGTCACAAGTATGGC 845

Qy         280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299
               |||||||||||||||:::||||||||||||||||||:::|||||||||||||||:::|||
Db         846 CTTGTATATGCTGGTGTCGGTTGGGTGATATGGCGGAGCAAGGAAGACTTGCCCGATGAA 905

Qy         300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
               |||:::|||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db         906 CTCGTCTTTCATATAAACTACCTTGGGTCTGATCAGCCTACTTTTACTCTCAACTTCTCT 965

Qy         320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
               |||||||||   |||:::|||||||||||||||||||||||||||||||||   ||||||
Db         966 AAAGGTTCCTATCAAATAATTGCACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGT 1025

Qy         340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
               |||:::|||||||||:::||||||      |||      ||||||   ||||||:::
Db        1026 TATAAGAACGTCATGAAGAATTGCTTATCAAACGCAAAAGTACTAACAGAGGGAATCACA 1085

Qy         360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
               |||      ||||||:::|||||||||||||||   |||||||||:::||||||||||||
Db        1086 AAAATGGGGCGGTTCGATATTGTCTCTAAGGATGTGGGTGTTCCTGTTGTAGCATTTTCT 1145

Qy         380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGly 399
               |||:::|||||||||   :::|||   ||||||:::|||:::   |||||||||:::|||
Db        1146 CTCAGGGACAGCAGCAAATATACGGTATTTGAAGTATCTGAGCATCTCAGAAGATTTGGA 1205

Qy         400 TrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgVal 419
               ||||||||||||||||||||||||||||||:::|||:::||||||   ||||||||||||
Db        1206 TGGATCGTCCCTGCATACACAATGCCACCGGATGCTGAACACATTGCTGTACTGCGGGTT 1265

Qy         420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
               |||||||||||||||||||||   :::||||||||||||||||||   ||||||||||||
Db        1266 GTCATTAGAGAGGATTTCAGCCACAGCCTAGCTGAGAGACTTGTTTCTGACATTGAGAAA 1325

Qy         440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIle------SerLeu 457
               ::::::   |||||||||      |||   |||:::      |||
Db        1326 ATTCTGTCAGAGTTGGACACACAGCCTCCTCGTTTGCCCACCAAAGCTGTCCGTGTCACT 1385

Qy         458 GlyGlnGluLysSerGluSerAsnSerAspAsnLeu------------MetValThrVal 473
                  :::|||      :::         |||   |||            |||   ||||||
Db        1386 GCTGAGGAAGTGCGTGATGACAAGGGTGATGGGCTTCATCATTTTCACATGGATACTGTA 1445

Qy         474 LysLysSerAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAla 493
                              :::   |||:::|||||||||      |||:::|||
Db        1446 ---------------GAGACTCAGAAAGACATTATCAAACATTGGAGGAAAATCGCAGGG 1490
```

```
y        494 AspArgLysLysThrSerGlyIleCys 502
                   |||||||||||||||:::|||
b       1491 ------AAGAAGACCAGCGGAGTCTGC 1511

ESULT 2
S-09-103-840A-2/c
 Sequence 2, Application US/09103840A
 Patent No. 6294328
 GENERAL INFORMATION:
  APPLICANT: FLEISCHMAN, Robert D.
  APPLICANT:  WHITE, Owen R.
  APPLICANT:  FRASER, Claire M.
  APPLICANT:  VENTER, John C.
  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
  TITLE OF INVENTION:  TUBERCULOSIS
  FILE REFERENCE: 24366-20007.00
  CURRENT APPLICATION NUMBER: US/09/103,840A
  CURRENT FILING DATE:  1998-06-24
  NUMBER OF SEQ ID NOS: 2
  SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2
   LENGTH: 4403765
   TYPE: DNA
   ORGANISM: Mycobacterium tuberculosis
   FEATURE:
   OTHER INFORMATION: CDC 1551
   OTHER INFORMATION: "n" bases at various positions throughout the sequence
   OTHER INFORMATION: represent a, t, c or g
S-09-103-840A-2

lignment Scores:
red. No.:              6.33e-122      Length:          4403765
core:                  1114.00        Matches:         219
ercent Similarity:     65.63%         Conservative:    77
est Local Similarity:  48.56%         Mismatches:      143
uery Match:            42.60%         Indels:          12
B:                     3              Gaps:            5

S-10-006-852-2 (1-502) x US-09-103-840A-2 (1-4403765)

y          6 AlaValSerGluSerAspValSerVal------HisSer---------ThrPheAlaSer 20
             |||||||||   |||      ||||||      ||||||            :::
b    3837651 GCCGTGTCTCGCAGTCACCCGTCCGTGCCCGCGCATTCGATCGCTCCGGCCTACACCGGT 3837592

y         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
             |||               :::|||      :::||||||:::   |||:::      |||
b    3837591 CGGATGTTCACCGCACCGGTGCCGGCGCTGCGGATGCCCGATGAGTCGATGGATCCCGAG 3837532

y         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
             |||||||||:::   |||:::|||||||||||||||||||||:::   ||||||||||||
b    3837531 GCGGCCTATCGCTTCATCCACGACGAGCTGATGCTCGACGGTAGCTCGCGGCTGAATCTG 3837472

Qy        61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
             |||:::||||||||||||||||||:::||||||   :::||||||:::      :::
Db   3837471 GCCACCTTCGTGACCACCTGGATGGACCCGGAGGCCGAAAAGCTGATGGCCGAGACGTTC 3837412

Qy        81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
             :::||||||   :::|||   ||||||||||||   |||      ::::::   ||||||
Db   3837411 GACAAGAACATGATCGATAAGGACGAATACCCGGCGACCGCGGCCATCGAGGCGCGCTGT 3837352

Qy       101 ValAsnMetIleAlaHisLeuPheAsnAlaPro---------LeuGluGluAlaGluThr 117
             |||:::|||:::|||   ||||||:::|||                  :::      :::
Db   3837351 GTGTCCATGGTCGCCGACCTGTTTCACGCCGAGGGTCTGCGCGACCACGACCCCACCAGC 3837292

Qy       118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137
             |||   ||||||   |||:::|||||||||||||||:::||||||   |||||||||
Db   3837291 GCCACCGGGGTGTCCACCATCGGCTCCAGCGAGGCGGTGATGCTGGGTGGGCTGGCCCTG 3837232

Qy       138 LysArgLysTrpGlnAsnLysArgLysAla---GluGlyLysProValAspLysProAsn 156
             |||   :::|||:::   :::      :::   :::|||:::            ||||||
Db   3837231 AAATGGCGTTGGCGGCAACGGGTGGGGTCTTGGAAGGGGCGC---------ATGCCCAAT 3837181

Qy       157 IleValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluVal 176
             :::|||   |||:::||||||||||||   ||||||||||||   |||||||||:::|||
Db   3837180 CTGGTGATGGGTTCGAACGTCCAGGTGGTGTGGGAGAAGTTCTGCCGCTACTTCGACGTC 3837121

Qy       177 GluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaVal 196
             |||   :::   :::   :::      |||   ||||||:::   |||:::|||   :::
Db   3837120 GAACCCCGTTACCTGCCGATGGAGCGGGGCCGCTACGTCATCACCCCCGAGCAGGTGCTC 3837061

Qy       197 AspMetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGly 216
                   ||||||||||||||||||   |||      |||||||||:::|||      |||
Db   3837060 GCCGCCGTCGACGAGAACACCATCGGCGTGGTGGCGATCTTGGGCACCACCTATACCGGT 3837001

Qy       217 GluPheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGly 236
             |||   |||   :::      :::         |||   :::               |||
Db   3837000 GAACTCGAACCCATCGCCGAGATCTGCGCCGCGCTGGACAAACTGGCGGCTGGCGGGGT 3836941

Qy       237 TrpAspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyr 256
                |||   |||:::|||||||||||||||||||||||||||:::   |||||||||:::
Db   3836940 GTGGACGTCCCGGTACACGTCGACGCGGCCAGTGGGGGCTTTGTGGTGCCGTTTTTGCAT 3836881

Qy       257 ProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHis 276
             |||:::|||   ||||||||||||||||||   |||   |||||||||||||||||||||
Db   3836880 CCGGACCTGGTATGGGATTTTCGGCTGCCCCGCGTGGTGTCGATCAACGTCAGCGGCCAC 3836821

Qy       277 LysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeu 296
             ||||||||||||   |||   |||:::|||:::|||:::||||||      |||   |||
Db   3836820 AAGTATGGGCTGACCTATCCCGGCGTCGGGTTTGTCGTGTGGCGCGGGCCCGAGCACCTG 3836761

Qy       297 ProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeu 316
             ||||||:::|||:::|||   :::||||||||||||   |||   |||||||||||||||
Db   3836760 CCGGAGGATCTGGTTTTCCGGGTCAACTACCTCGGCGGCGACATGCCGACCTTCACCCTG 3836701

Qy       317 AsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGly 336
             |||||||||:::      :::||||||:::   |||||||||      :::|||||||||
Db   3836700 AACTTCTCCCGTCCCGGTAACCAGGTGGTGGCCCAGTACTACAACTTCCTGCGGCTGGGG 3836641

Qy       337 HisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
                :::||||||      ||||||:::                        |||   :::
Db   3836640 CGCGACGGCTATACCAAGGTGATGCAGGCGCTGTCGCACACCGCCCGGTGGCTGGGTGAC 3836581

Qy       357 GlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuVal 376
                |||   :::   :::         ::::::|||            :::|||:::|||
Db   3836580 CAGCTGCGCGAGGTGGATCATTGCGAGGTGATCTCGGATGGTTCGGCGATCCCGGTGGTC 3836521

Qy       377 AlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArg 396
             :::|||   |||               :::|||||||||::::::|||      ||||||
Db   3836520 AGCTTCCGGCTCGCCGGCGACCGCGGGTACACGGAGTTCGACGTCTCCCATGAGCTGCGG 3836461

Qy       397 ArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrVal 416
                :::||||||   |||||||||||||||||||||   ||||||      :::   |||
Db   3836460 ACCTTCGGGTGGCAGGTGCCCGCCTACACCATGCCGGACAACGCCACCGACGTGGCGGTG 3836401

Qy       417 LeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAsp 436
             ||||||:::|||:::||||||      |||      ||||||      |||      |||
Db   3836400 CTGCGGATCGTGGTTCGGGAAGGACTCTCCGCCGACCTGGCGCGGGCCCTGCACGACGAC 3836341

Qy       437 IleGluLysValMetArgGluLeuAspGluLeu 447
                         :::      ||||||::::::
Db   3836340 GCGGTCACCGCGTTGGCTGCCCTGGACAAGGTC 3836308

RESULT 3
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
```

```
  CURRENT FILING DATE:  1998-06-24
  NUMBER OF SEQ ID NOS: 2
  SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
   LENGTH: 4411529
   TYPE: DNA
   ORGANISM: Mycobacterium tuberculosis
   OTHER INFORMATION: H37Rv
S-09-103-840A-1

lignment Scores:
red. No.:              6.35e-122      Length:        4411529
core:                  1114.00        Matches:       219
ercent Similarity:     65.63%         Conservative:  77
est Local Similarity:  48.56%         Mismatches:    143
uery Match:            42.60%         Indels:        12
B:                     3              Gaps:          5

S-10-006-852-2 (1-502) x US-09-103-840A-1 (1-4411529)

y          6 AlaValSerGluSerAspValSerVal------HisSer---------ThrPheAlaSer 20
             |||||||||   |||      ||||||      ||||||            :::
b    3851753 GCCGTGTCTCGCAGTCACCCGTCCGTGCCCGCGCATTCGATCGCTCCGGCCTACACCGGT 3851694

y         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
             |||               :::|||      :::||||||:::   |||:::      |||
b    3851693 CGGATGTTCACCGCACCGGTGCCGGCGCTGCGGATGCCCGATGAGTCGATGGATCCCGAG 3851634

y         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
             |||||||||:::   |||:::|||||||||||||||||||||:::   ||||||||||||
b    3851633 GCGGCCTATCGCTTCATCCACGACGAGCTGATGCTCGACGGTAGCTCGCGGCTGAATCTG 3851574

y         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
             |||:::||||||||||||||||||:::||||||   :::||||||:::      :::
b    3851573 GCCACCTTCGTGACCACCTGGATGGACCCGGAGGCCGAAAAGCTGATGGCCGAGACGTTC 3851514

y         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
             :::||||||   :::|||   ||||||||||||   |||      ::::::   ||||||
b    3851513 GACAAGAACATGATCGATAAGGACGAATACCCGGCGACCGCGGCCATCGAGGCGCGCTGT 3851454

y        101 ValAsnMetIleAlaHisLeuPheAsnAlaPro---------LeuGluGluAlaGluThr 117
             |||:::|||:::|||   ||||||:::|||                  :::      :::
b    3851453 GTGTCCATGGTCGCCGACCTGTTTCACGCCGAGGGTCTGCGCGACCACGACCCCACCAGC 3851394

y        118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137
             |||   ||||||   |||:::|||||||||||||||:::||||||   |||||||||
b    3851393 GCCACCGGGGTGTCCACCATCGGCTCCAGCGAGGCGGTGATGCTGGGTGGGCTGGCCCTG 3851334

y        138 LysArgLysTrpGlnAsnLysArgLysAla---GluGlyLysProValAspLysProAsn 156
             |||   :::|||:::   :::      :::   :::|||:::         ||||||
b    3851333 AAATGGCGTTGGCGGCAACGGGTGGGGTCTTGGAAGGGGCGC---------ATGCCCAAT 3851283

y        157 IleValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluVal 176
             :::|||   |||:::||||||||||||   ||||||||||||   |||||||||:::|||
b    3851282 CTGGTGATGGGTTCGAACGTCCAGGTGGTGTGGGAGAAGTTCTGCCGCTACTTCGACGTC 3851223

Qy       177 GluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaVal 196
             |||   :::   :::   :::      |||   ||||||:::   |||:::|||   :::
Db   3851222 GAACCCCGTTACCTGCCGATGGAGCGGGGCCGCTACGTCATCACCCCCGAGCAGGTGCTC 3851163

Qy       197 AspMetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGly 216
                   ||||||||||||||||||   |||      |||||||||:::|||      |||
Db   3851162 GCCGCCGTCGACGAGAACACCATCGGCGTGGTGGCGATCTTGGGCACCACCTATACCGGT 3851103

Qy       217 GluPheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGly 236
             |||   |||   :::      :::         |||   :::               |||
Db   3851102 GAACTCGAACCCATCGCCGAGATCTGCGCCGCGCTGGACAAACTGGCGGCTGGCGGGGGT 3851043

Qy       237 TrpAspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyr 256
                |||   |||:::|||||||||||||||||||||||||||:::   |||||||||:::
Db   3851042 GTGGACGTCCCGGTACACGTCGACGCGGCCAGTGGGGGCTTTGTGGTGCCGTTTTTGCAT 3850983

Qy       257 ProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHis 276
```

```
             |||:::|||   ||||||||||||||||||   |||   |||||||||||||||||||||
Db   3850982 CCGGACCTGGTATGGGATTTTCGGCTGCCCCGCGTGGTGTCGATCAACGTCAGCGGCCAC 3850923

Qy       277 LysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeu 296
             ||||||||||||   |||   |||:::|||:::|||:::||||||      |||   |||
Db   3850922 AAGTATGGGCTGACCTATCCCGGCGTCGGGTTTGTCGTGTGGCGCGGGCCCGAGCACCTG 3850863

Qy       297 ProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeu 316
             ||||||:::|||:::|||   :::||||||||||||   |||   |||||||||||||||
Db   3850862 CCGGAGGATCTGGTTTTCCGGGTCAACTACCTCGGCGGCGACATGCCGACCTTCACCCTG 3850803

Qy       317 AsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGly 336
             |||||||||:::      :::||||||:::   |||||||||      :::|||||||||
Db   3850802 AACTTCTCCCGTCCCGGTAACCAGGTGGTGGCCAGTACTACAACTTCCTGCGGCTGGGG 3850743

Qy       337 HisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
                :::||||||      ||||||:::                        |||   :::
Db   3850742 CGCGACGGCTATACCAAGGTGATGCAGGCGCTGTCGCACACCGCCCGGTGGCTGGGTGAC 3850683

Qy       357 GlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuVal 376
                |||   :::   :::         ::::::|||            :::|||:::|||
Db   3850682 CAGCTGCGCGAGGTGGATCATTGCGAGGTGATCTCGGATGGTTCGGCGATCCCGGTGGTC 3850623

Qy       377 AlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArg 396
             :::|||   |||               :::|||||||||::::::|||      ||||||
Db   3850622 AGCTTCCGGCTCGCCGGCGACCGCGGGTACACGGAGTTCGACGTCTCCCATGAGCTGCGG 3850563

Qy       397 ArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrVal 416
                :::||||||   |||||||||||||||||||||   ||||||      :::   |||
Db   3850562 ACCTTCGGGTGGCAGGTGCCCGCCTACACCATGCCGGACAACGCCACCGACGTGGCGGTG 3850503

Qy       417 LeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAsp 436
             ||||||:::|||:::||||||      |||      ||||||      |||      |||
Db   3850502 CTGCGGATCGTGGTTCGGGAAGGACTCTCCGCCGACCTGGCGCGGGCCCTGCACGACGAC 3850443

Qy       437 IleGluLysValMetArgGluLeuAspGluLeu 447
                         :::      ||||||::::::
Db   3850442 GCGGTCACCGCGTTGGCTGCCCTGGACAAGGTC 3850410

RESULT 4
US-09-068-195-21
; Sequence 21, Application US/09068195B
; Patent No. 6140078
; GENERAL INFORMATION:
;  APPLICANT: Sanders, Jan W.
;  APPLICANT: Ledeboer, Adrianus M.
;  APPLICANT: Venema, Gerard
;  APPLICANT: Kok, Jan
;  TITLE OF INVENTION: Salt-Inducible Promoter Derivable from a Lactic Acid
;  TITLE OF INVENTION: Bacterium, and Its Use in a Lacctic Acid Bacterium for
;  TITLE OF INVENTION: Production of Desired Protein
;  FILE REFERENCE: Sanders-60113/0252227
;  CURRENT APPLICATION NUMBER: US/09/068,195B
;  CURRENT FILING DATE: 1998-07-29
;  EARLIER APPLICATION NUMBER: PCT/EP97/04755
;  EARLIER FILING DATE: 1997-08-20
;  EARLIER APPLICATION NUMBER: EP 97200744/7
;  EARLIER FILING DATE: 1997-03-13
;  EARLIER APPLICATION NUMBER: EP 96202444/4
;  EARLIER FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 5565
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: genomic DNA of
;   OTHER INFORMATION: NS3 locus containing gadRCB
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (3)..(500)
;   OTHER INFORMATION: C-terminus of rnhB
```

```
   FEATURE:
   NAME/KEY: mat_peptide
   LOCATION: (1095)..(1922)
   OTHER INFORMATION: rggL = gadR
   FEATURE:
   NAME/KEY: mat_peptide
   LOCATION: (2069)..(3577)
   OTHER INFORMATION: orfX = gadC
   FEATURE:
   NAME/KEY: mat_peptide
   LOCATION: (3600)..(4997)
   OTHER INFORMATION: gadB
   FEATURE:
   NAME/KEY: mat_peptide
   LOCATION: Complement((5078)..(5563))
   OTHER INFORMATION: C-terminus of ORF in opposite direction
S-09-068-195-21

lignment Scores:
red. No.:              8.5e-116      Length:        5565
core:                  1025.00       Matches:       203
ercent Similarity:     63.46%        Conservative:  94
est Local Similarity:  43.38%        Mismatches:    156
uery Match:            39.20%        Indels:        15
B:                     3             Gaps:          4

S-10-006-852-2 (1-502) x US-09-068-195-21 (1-5565)

y          18 PheAlaSerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIle 37
              |||   |||         :::      ||||||::::::|||:::   :::   ||||||
b        3651 TTTGGTTCAGAAAGTGAACAAGTGGATTTACCTAAATATAAATTAGCTCAACAATCAATT 3710

y          38 ProLysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArg 57
                          |||||||||::::::   ||||||::::::   :::||||||   |||
b        3711 GAGCCTCGAGTGGCCTATCAGTTAGTTCAAGATGAAATGCTAGATGAAGGAACGCTCGT 3770

y          58 LeuAsnLeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMet 77
              ||||||||||||:::|||      |||:::||||||||||||      ||||||:::
b        3771 TTAAATTTGGCCACATTCTGTCAAACTTATATGGAACCTGAAGCAGTCAAGCTGATGAGT 3830

y          78 SerSerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGln 97
                 ::::::   ||||||   :::|||      |||||||||   |||||||||::::::
b        3831 CAGACCTTGGAAAAAAATGCGATTGACAAATCAGAATATCCAAGAACAACTGAAATTGAA 3890

y          98 AsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThr 117
              ||||||||||||||||||||||||   |||:::||||||         |||
b        3891 AACCGTTGCGTCAACATGATCGCTGACCTTTGGAATGCG------AGTGAAAAAGGAAAA 3944

y         118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137
                    |||      |||:::|||||||||||||||   ||||||   |||:::|||
b        3945 ATTTATGGGACTTCGACAATTGGTTCTTCAGAAGCTTGTATGCTTGGGGGAATGGCTATG 4004

y         138 LysArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAsp------LysPro 155
              |||      |||:::   :::   :::      |||      ::::::      ||||||
b        4005 AAGTTTTCTTGGCGTAAGCGAGCAGAAAAATTAGGCCTAGATATTAATGCGAAAAAGCCA 4064

y         156 AsnIleValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGlu 175
              |||:::|||      :::      |||||||||||||||||||||      |||::::::
b        4065 AACTTAGTCATTTCCTCTGGTTATCAAGTTTGCTGGGAAAAATTCTGTGTTTATTGGGAT 4124

y         176 ValGluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAla 195
              :::|||::::::||||||   :::         :::      ::::::   ::::::
b        4125 ATTGAAATGAGAGAAGTGCCAATGGATAGAGAACATATGTCAATCAATTTGGAAAAAGTG 4184

y         196 ValAspMetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsn 215
              :::|||   |||||||||   ||||||   |||      |||:::|||   |||
b        4185 ATGGATTATGTTGATGAATATACGATTGGAGTAGTTGGAATTATGGGGATTACTTATACT 4244

y         216 GlyGluPheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThr 235
              |||   ::::::|||:::|||   |||::::::|||:::   |||   ||||||:::|||
b        4245 GGTCGTTATGATGATATCAAAGCTTTGGATAATTTGATTGAAGAATATAATAAACAGACA 4304

y         236 GlyTrpAspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu 255
                 :::         |||||||||||||||||||||||||||      |||||||||:::
Db       4305 GACTACAAAGTTTATATTCACGTAGATGCTGCTTCAGGAGGACTTTATGCTCCTTTTGTT 4364

Qy        256 TyrProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGly 275
                 |||||||||||||||||||||||||||      |||   |||||||||   ||||||
Db       4365 GAGCCAGAACTTGAGTGGGATTTCCGTTTGAAAAATGTCATTTCAATCAATACTTCAGGA 4424

Qy        276 HisLysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAsp 295
              |||||||||||||||||||||   |||:::|||||||||:::||||||:::|||:::
Db       4425 CATAAATATGGTTTAGTATATCCTGGTGTAGGTTGGGTCTTGTGGCGTGACAAAAAATAT 4484

Qy        296 LeuProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThr 315
              |||||||||||||||||||||   ::::::|||||||||   :::   ||||||
Db       4485 TTACCTGAAGAGTTAATTTTTAAAGTAAGTTATCTTGGAGGAGAATTACCAACAATGGCG 4544

Qy        316 LeuAsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeu 335
              :::|||||||||      :::||||||:::|||   |||||||||      :::|||
Db       4545 ATTAATTTTTCTCACAGTGCTTCTCAATTAATCGGTCAATACTATAATTTTGTACGTTAT 4604

Qy        336 GlyHisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArg 355
              |||   :::||||||:::   :::   |||         :::      :::   |||
Db       4605 GGATTTGATGGATATAAAGCTATTCATGAGAGAACGCATAAAGTAGCCATGTATTTAGCA 4664

Qy        356 GluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeu 375
              |||   :::|||||||||      |||   |||::::::            :::|||:::
Db       4665 GAAGAAATTGAAAAAACAGGAATGTTTGAGATTATGAACGATGGGGCACAATTACCAATT 4724

Qy        376 ValAlaPheSerLeuLysAspSerSerCysHis------ThrGluPheGluIleSerAsp 393
              |||   :::   ||||||::::::|||                  ::::::::::::|||
Db       4725 GTCTGCTACAAATTAAAAGAAAATTCAAACCGTGGTTGGAATCTTTATGATTTGGCAGAT 4784

Qy        394 MetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHis 413
                 |||         ||||||   ||||||||||||   :::|||   |||   ::::::
Db       4785 CGTTTATTAATGAAGGGATGGCAAGTGCCTGCTTATCCACTTCCTAAAAATTTGGAAAAT 4844

Qy        414 IleThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeu 433
                    :::   |||:::|||||||||   ||||||         :::|||
Db       4845 GAAATCATTCAACGTTTAGTAATTCGAGCAGATTTCGGGATGAATATGGCATTTAACTAT 4904

Qy        434 ValIleAspIleGluLysValMetArgGluLeuAspGlu-LeuProSerArgValIleHi 453
              |||   |||:::::::::   :::      |||:::::: |||            |||
Db       4905 GTTCAAGATATGCAAGAAGCAATTGATGCACTAAATAAGGCTCATATTCTATTTCATCAG 4964

Qy        453 sLysIleSerLeuGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrVa 473
                  :::   :::                                    ::::::|||
Db       4965 GAACCTGAAAATA-----------------------AAACATATGGCTTTACTCACTAA 5000

Qy        473 lLysLysSerAspIleAspLys 480
                  ||||||||||||   |||
Db       5001 AGATAAAAGCGATATATCTAAG 5022

RESULT 5
US-08-311-731A-130/c
; Sequence 130, Application US/08311731A
; Patent No. 6583266
;  GENERAL INFORMATION:
;    APPLICANT:  SMITH, DOUGLAS
;    APPLICANT:  MAO, JEN-I
;    TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES
;    TITLE OF INVENTION:  RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
;    TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;    NUMBER OF SEQUENCES:  411
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WOLF, GREENFIELD & SACKS, P.C.
;      STREET:  600 ATLANTIC AVENUE
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/311,731A
      FILING DATE:
      CLASSIFICATION:  530
    ATTORNEY/AGENT INFORMATION:
      NAME:  GATES, EDWARD R.
      REGISTRATION NUMBER:  31,616
      REFERENCE/DOCKET NUMBER:  C0044/7125
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  617/720-3500
      TELEFAX:  617/720-2441
  INFORMATION FOR SEQ ID NO:  130:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  36941 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  circular
    MOLECULE TYPE:  DNA (genomic)
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    ORIGINAL SOURCE:
      ORGANISM:  MYCOBACTERIUM LEPRAE
S-08-311-731A-130

lignment Scores:
red. No.:              2.27e-62       Length:        36941
core:                  605.00         Matches:       175
Percent Similarity:    55.86%         Conservative:  68
Best Local Similarity: 40.23%         Mismatches:    130
Query Match:           23.14%         Indels:        67
DB:                    4              Gaps:          7

US-10-006-852-2 (1-502) x US-08-311-731A-130 (1-36941)

Qy         32 MetProGluAsnSerIleProLysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMet 51
              :::||||||   |||:::      ||||||||||||:::   |||:::|||   ||||||
Db      10475 CTGCCCGAAGCATCGATGGATCCTGAGGCGGCGTACCGCTTTATTCACGACTAGCTAATG 10416

Qy         52 LeuAspGlyAsnProArgLeuAsnLeu--AlaSerPheValThrThrTrpMetGluProG 71
              :::||||||:::   ||||||||||||  |||:::|||||||||||||||::::::||||
Db      10415 GTAGACGGCAGTTCTCGGCTCAACCTAGCGCTACTTTCGTCACCACCTGGGTGGATCCCG 10356

Qy         71 luCysAspLysLeuIleMetSerSerIleAsnLysAsnTyrValAspMetAspGluTyrP 91
              ||      ||||||:::      :::   :::|||      :::|||   ||||||||||
Db      10355 AGGTAAAGAAGTTGATGGCCGAGACGTTCGACAAAACATGTTTAGACAAGGACGAGTATC 10296

Qy         91 roValThrThrGluLeuGlnAsnArgCysVal-AsnMetIleAlaHisLeuPheAsnAla 110
              ||            ::::::   ||||||||| :::|||         |||::::::
Db      10295 CGGCGGCCGCGCC-ATCGAACGGCGCTGCGTATAGCATGGCGACCGGTCTCTATTCACG- 10238

Qy        111 ProLeuGluGluAlaGluThrAla------------ValGlyValGlyThrValGlySer 126
              |||   :::   |||   |||               ||||||         :::||||||
Db      10237 CCTAGGGACTCAGCTACGACAACCCGTCCAATACCTGTGGGGTGTTCCATTATCGGCTCC 10178

Qy        127 SerGlu--AlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrp------------- 141
                 :::     :::||||||      ||||||   |||   :::|||
Db      10177 AGGGATGGGTGGTGATGCTGGGTTGCCTGGCCATAAAATTGCGCTGGTGTGCGGTAGCTA 10118

Qy        142 -----GlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrG 160
                   |||                                    ||||||:::      |
Db      10117 GTTAGCAGCGT---------------------------AGCACCCCCAATCTGGGGATGG 10085

Qy        160 lyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysG 180
              |||||   |||||||||   ||||||||||||   |||||||||:::||||||   :::
Db      10084 GCGCAGGGGTCCAGGTGGTTTGGGAGAAGTTCTGTCGTTACTTCGATGTCGAGC--CGCT 10027

Qy        180 luValLysLeuSerGluGlyTyrTyrValMet-AspProGlnGlnAlaValAspMet-Va 199
                :::   :::   ||||||   ||||||:::    |||:::|||   ::::::    ||
Db      10026 ATCTGTCGATGGAGGAAGGATGCTACGTCATCACCTCCCGAGCAGGTAATCGAGGCCCGT 9967

Qy        199 lAspGluAsnThrIle---CysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPh 218
              |||||||:::|||      |||            |||   :::|||      ||||||
Db       9966 CGACGAGGACACTACTTGGTGTCGTGGCCATCT-TTGTCTGCCACTTTCACCGGCGAACT 9908

Qy        218 eGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAs 238
               |||   |||      |||         |||   :::         :::   |||   ||
Db       9907 CGAACCCGTCGCGGAGCTCTGCGTCGCGCTGGACAAGCTGGCAGCCGACAATGGCTTGGA 9848

Qy        238 pThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGl 258
              |      :::   |||   |||||||||||||||        |||    |||:::|||||
Db       9847 CATTGCAGTGTCCGTCGGTGCCGCCAGCGGGGGTTCGGG-ATGCCTTGCTTGCACCCGGA 9789

Qy        258 uLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnVal-SerGlyHisLysT 278
              ||||:::||||||      ||||||   |||   |||||||||||| ||||||:::||||
Db       9788 GTTGAAGTGGGATTCCTGGCTGCCCCGAGTGATGTCGATCAATGTAAAGTGGCTATAAAT 9729

Qy        278 yrGlyLeuVal-TyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuPro 297
              ||||||||    |||   |||:::|||:::   :::||||||      |||   ||||||
Db       9728 ATGGGCTCACCCTACCCCGGTGTGGGATTTGGCGTATGGCGCGGCAACGAATATTTGCCC 9669

Qy        298 GluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsn 317
              |||:::::::::|||   :::||||||||||||   |||      ||||||   ||||||
Db       9668 GAGGACGTAGTCTTTCGTGTCAACTACCTGGGCGGCGATATGTAGACTTTCGTATTGAAC 9609

Qy        318 PheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHis 337
              ||||||:::      :::||||||::::::|||||||||      :::   ||||||
Db       9608 TTTTCCCGCCTGGACAACCAAGTGGTTAGTCAGTACTACAACTTTCTGTGGCTGGGACGT 9549

Qy        338 GluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu--- 356
                 ||||||      ||||||:::               :::      |||   |||
Db       9548 AACGGTTACTGGAAAGTCATGCAGGCCTTGTCGCTGACGGTGCGGTGGCTGATTGAGTAG 9489

Qy        357 ---------------------------------------GlyLeuGluLysThrGluArg 363
                                                     ||||||      |||   |||
Db       9488 CTCCGCGTCGGTGACTATTGTGAGCTGATCGGTTACATCGGACTATCAGACACCGCACGG 9429

Qy        364 ---------------PheAsnIleValSerLysAspGluGlyValProLeuValAlaPhe 378
                                ::::::::::   :::|||   |||
Db       9428 CCATCCTGGTGGTATACCTCCTTGCTCGTTAGAGACTGTGGG------------------ 9387

Qy        379 SerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyr 398
                                   :::|||||||||::::::         ||||||   :::
Db       9386 ---------------------TACACCGAGTTCGATCTCCTCCACGAACTCCGGACATTC 9348

Qy        399 GlyTrpIleValProAlaTyr-ThrMetProProAsnAlaGlnHisIleThrValLeuAr 418
              ||||||   |||||||||||| |||:::|||   :::|||      ::::::||||||||
Db       9347 GGTTGGCAGGTGCCCGCCTATAACTGTGCCCGACGATGCCACTGACGTATCGGTGCTGCG 9288

Qy        418 gValValIleArgGluAspPheSer 426
              |:::|||:::   |||      |||
Db       9287 CATCGTGGTCTGGGAAGGGCTCTCG 9263

RESULT 6
US-08-939-309-7
; Sequence 7, Application US/08939309
; Patent No. 6423527
;  GENERAL INFORMATION:
;    APPLICANT:  Saba, Julie D.
;    APPLICANT:  Zhou, Jianhui
;    TITLE OF INVENTION:  SPHINGOSINE-1-PHOSPHATE LYASE
;    TITLE OF INVENTION:  POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
;    TITLE OF INVENTION:  METHODS OF USE THEREFOR
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/939,309
        FILING DATE:  29-SEP-1997
        CLASSIFICATION:  800
    ATTORNEY/AGENT INFORMATION:
      NAME:  David, Maki J.
      REGISTRATION NUMBER:  31,392
      REFERENCE/DOCKET NUMBER:  200116.402
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (206) 622-4900
      TELEFAX:  (206) 682-6031
  INFORMATION FOR SEQ ID NO:  7:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1770 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  1..1767
S-08-939-309-7

lignment Scores:
red. No.:               6.54e-30       Length:        1770
core:                   325.00         Matches:       122
ercent Similarity:      44.53%         Conservative:  98
est Local Similarity:   24.70%         Mismatches:    188
uery Match:             12.43%         Indels:        86
B:                      4              Gaps:          20

S-10-006-852-2 (1-502) x US-08-939-309-7 (1-1770)

y         3 LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22
            |||         |||   :::            |||      :::
b       316 TTGAGGGGTACCGTAGAAAAGGAAGTCACAAAGGTCAAACAATCGATCGAAGACGAACTA 375

y        23 ValArg------ThrSerLeuProArgPhe---LysMetProGluAsnSerIleProLys 39
            :::|||      :::   |||      |||   ::::::|||   |||   ||||||:::
b       376 ATTAGATCGGACTCTCAGTTAATGAATTTCCCACAGTTGCCATCCAATGGGATACCTCAG 435

y        40 GluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
            :::         :::   :::|||                           :::||||||
b       436 GATGATGTTATTGAAGAGCTAAAT---------------------------AAATTGAAC 468

y        60 LeuAlaSerPheValThrThrTrpMetGluProGluCys--------------------- 72
                           |||   |||   |||   :::
b       469 GACTTGATACCACATACCCAATGGAAGGAAGGAAAGGTCTCTGGTGCCGTTTACCACGGT 528

y        73 ------AspLysLeuIleMetSerSerIleAsn------------------LysAsnTyr 84
                  |||   :::   :::   :::|||                        |||
b       529 GGTGATGATTTGATCCACTTACAAACAATCGCATACGAAAAATATTGCGTTGCCAATCAA 588

y        85 ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle 104
            :::      |||   :::|||         ::::::::::::      |||:::|||:::
b       589 TTACATCCCGATGTCTTTCCTGCCGTACGTAAAATGGAATCCGAAGTGGTTTCTATGGTT 648

y       105 AlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyValGlyThrVal 124
                  :::||||||||||||         ::::::|||      |||      |||
b       649 TTAAGAATGTTTAATGCCCCT---------TCTGATACAGGTTGTGGTACCACAACTTCA 699

Qy      125 GlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLys 144
            |||   :::|||:::::::::||||||   |||:::   |||            :::
Db      700 GGTGGTACAGAATCCTTGCTTTTAGCATGTCTGAGCGCTAAAATGTATGCCCTTCATCAT 759

Qy      145 ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln 164
            |||                  :::   :::|||   |||:::
Db      760 CGT---------------GGAATCACCGAACCAGAAATAATTGCTCCCGTAACTGCACAT 804

Qy      165 ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer 184
                  ::::::|||   |||   ||||||   ::::::|||:::   |||:::|||
Db      805 GCTGGGTTTGACAAAGCTGCTTATTACTTTGGCATGAAGCTACGCCACGTGGAGCTAGAT 864

Qy      185 GluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle 204
                     |||   :::|||      :::            :::::::::|||||||||
Db      865 CCAACGACATATCAAGTGGACCTGGGAAAAGTGAAAAAATTCATCAATAAGAACACAATT 924

Qy      205 CysValAlaAspIleLeuGlySerThrLeuAsn---------GlyGluPheGluAspVal 221
                        ::::::||||||      |||         |||      :::|||:::
Db      925 TTA---------CTGGTCGGTTCCGCTCCAAACTTTCCTCATGGTATTGCCGATGATATT 975

Qy      222 LysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle 241
            :::   |||         :::   :::|||   |||                  |||:::
Db      976 GAAGGATTGGGTAAA---ATAGCACAAAAATATAAA---------------CTTCCTTTA 1017

Qy      242 HisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu------------TyrPro 257
            |||||||||:::      |||   ||||||      |||:::            |||
Db     1018 CACGTCGACAGTTGTCTAGGTTCCTTTATTGTTTCATTTATGGAAAAGGCTGGTTACAAA 1077

Qy      258 GluLeuGlu---TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHis 276
               |||         |||||||||:::|||   |||   ||||||:::         |||
Db     1078 AATCTGCCATTACTTGACTTTAGAGTCCCGGGAGTCACCTCAATATCATGTGACACTCAT 1137

Qy      277 LysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeu 296
            |||||||||            |||         :::::::::||||||   :::
Db     1138 AAATATGGATTTGCACCAAAAGGCTCGTCAGTTATAATGTATAGAAACAGCGACTTACGA 1197

Qy      297 ProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeu 316
                  :::         :::      :::   |||            :::   ||||||
Db     1198 ATGCATCAGTATTACGTAAATCCTGCTTGGACTGGCGGGTTATATGGCTCTCCTACATTA 1257

Qy      317 AsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGly 336
                  |||:::      :::   |||:::      :::      ::::::   :::|||
Db     1258 GCAGGGTCCAGGCCTGGTGCTATTGTCGTAGGTTGTTGGGCCACTATGGTCAACATGGGT 1317

Qy      337 HisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
                  ||||||         :::|||:::|||:::|||   ::::::
Db     1318 GAAAATGGGTAC---------ATTGAGTCGTGCCAAGAA---ATAGTCGGTGCAGCAATG 1365

Qy      357 GlyLeuGluLysThrGluArgPheAsnIle------------ValSerLysAspGluGly 372
                  :::|||      :::   ||||||            :::      :::
Db     1366 AAGTTTAAAAAATACATCCAGGAAAACATTCCAGACCTGAATATAATGGGCAACCCTAGA 1425

Qy      373 ValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSer 392
                  :::::::::||||||   |||      :::   |||         |||:::|||
Db     1426 TATTCAGTCATTTCATTTTCTTCAAAGACCTTGAACATACAC---------GAACTATCT 1476

Qy      393 AspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGln 412
            |||   |||   :::   ||||||         |||         |||
Db     1477 GACAGGTTGTCCAAGAAAGGCTGGCATTTCAATGCCCTACAAAAGCCG------------ 1524

Qy      413 HisIleThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArg 432
                              |||   :::         |||:::|||   |||:::
Db     1525 ------------------GTTGCACTACACATGGCCTTCACGAGA---TTGAGCGCTCAT 1563

Qy      433 LeuValIleAspIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIle 452
            :::|||   :::|||      ::::::|||
Db     1564 GTTGTGGATGAGATCTGCGACATTTTACGT------------------------------ 1593

Qy      453 HisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSer 466
                     ::::::   :::   ||||||||||||||||||
Db     1594 ------ACTACCGTGCAAGAGTTGAAGAGCGAATCAAATTCT 1629

RESULT 7
US-09-849-180-7
; Sequence 7, Application US/09849180
; Patent No. 6495359
;   GENERAL INFORMATION:
;        APPLICANT: Saba, Julie D.
;                   Zhou, Jianhui
;        TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
;                            POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
;                            METHODS OF USE THEREFOR
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed Intellectual Property Law Group
;             STREET: 701 Fifth Avenue, Suite 6300
```

```
        CITY: Seattle
        STATE: Washington
        COUNTRY: USA
        ZIP: 98055
     COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: PatentIn Release #1.0, Version #1.30
     CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/09/849,180
        FILING DATE: 04-May-2001
        CLASSIFICATION: <Unknown>
     ATTORNEY/AGENT INFORMATION:
        NAME: Pepe, Jeffrey C.
        REGISTRATION NUMBER: 46,985
        REFERENCE/DOCKET NUMBER: 200116.402
     TELECOMMUNICATION INFORMATION:
        TELEPHONE: (206) 622-4900
        TELEFAX: (206) 682-6031
  INFORMATION FOR SEQ ID NO: 7:
     SEQUENCE CHARACTERISTICS:
        LENGTH: 1770 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: single
        TOPOLOGY: linear
     FEATURE:
        NAME/KEY:  CDS
        LOCATION:  1..1767
     SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-849-180-7

lignment Scores:
red. No.:              6.54e-30       Length:         1770
core:                  325.00         Matches:        122
ercent Similarity:     44.53%         Conservative:   98
est Local Similarity:  24.70%         Mismatches:     188
uery Match:            12.43%         Indels:         86
B:                     4              Gaps:           20

US-10-006-852-2 (1-502) x US-09-849-180-7 (1-1770)

y           3 LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22
              |||         |||   :::            |||      :::
b         316 TTGAGGGGTACCGTAGAAAAGGAAGTCACAAAGGTCAAACAATCGATCGAAGACGAACTA 375

y          23 ValArg------ThrSerLeuProArgPhe---LysMetProGluAsnSerIleProLys 39
              :::|||      :::   |||      |||   ::::::|||   |||   ||||||:::
b         376 ATTAGATCGGACTCTCAGTTAATGAATTTCCCACAGTTGCCATCCAATGGGATACCTCAG 435

y          40 GluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
              :::         :::   :::|||                           :::||||||
Db        436 GATGATGTTATTGAAGAGCTAAAT---------------------------AAATTGAAC 468

Qy         60 LeuAlaSerPheValThrThrTrpMetGluProGluCys--------------------- 72
                             |||   |||   |||   :::
b         469 GACTTGATACCACATACCCAATGGAAGGAAGGAAAGGTCTCTGGTGCCGTTTACCACGGT 528

Qy         73 ------AspLysLeuIleMetSerSerIleAsn------------------LysAsnTyr 84
                    |||   :::   :::   :::|||                        |||
Db        529 GGTGATGATTTGATCCACTTACAAACAATCGCATACGAAAAATATTGCGTTGCCAATCAA 588

Qy         85 ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle 104
              :::      |||   :::|||         ::::::::::::      |||:::|||:::
Db        589 TTACATCCCGATGTCTTTCCTGCCGTACGTAAAATGGAATCCGAAGTGGTTTCTATGGTT 648

Qy        105 AlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyValGlyThrVal 124
                    :::||||||||||||         ::::::|||      |||      |||
Db        649 TTAAGAATGTTTAATGCCCCT---------TCTGATACAGGTTGTGGTACCACAACTTCA 699

Qy        125 GlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLys 144
              |||   :::|||:::::::::||||||   |||:::   |||            :::
Db        700 GGTGGTACAGAATCCTTGCTTTTAGCATGTCTGAGCGCTAAAATGTATGCCCTTCATCAT 759

Qy        145 ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln 164
              |||                  :::   :::|||   |||:::
Db        760 CGT---------------GGAATCACCGAACCAGAAATAATTGCTCCCGTAACTGCACAT 804

Qy        165 ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer 184
                    ::::::|||   |||   ||||||   ::::::|||:::   |||:::|||
Db        805 GCTGGGTTTGACAAAGCTGCTTATTACTTTGGCATGAAGCTACGCCACGTGGAGCTAGAT 864

Qy        185 GluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle 204
                       |||   :::|||      :::            :::::::::|||||||||
Db        865 CCAACGACATATCAAGTGGACCTGGGAAAAGTGAAAAAATTCATCAATAAGAACACAATT 924

Qy        205 CysValAlaAspIleLeuGlySerThrLeuAsn---------GlyGluPheGluAspVal 221
                          ::::::||||||      |||         |||      :::|||:::
Db        925 TTA---------CTGGTCGGTTCCGCTCCAAACTTTCCTCATGGTATTGCCGATGATATT 975

Qy        222 LysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle 241
              :::   |||         :::   :::|||   |||                  |||:::
Db        976 GAAGGATTGGGTAAA---ATAGCACAAAAATATAAA---------------CTTCCTTTA 1017

Qy        242 HisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu------------TyrPro 257
              |||||||||:::      |||   ||||||      |||:::            |||
Db       1018 CACGTCGACAGTTGTCTAGGTTCCTTTATTGTTTCATTTATGGAAAAGGCTGGTTACAAA 1077

Qy        258 GluLeuGlu---TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHis 276
                 |||         |||||||||:::|||   |||   ||||||:::         |||
Db       1078 AATCTGCCATTACTTGACTTTAGAGTCCCGGGAGTCACCTCAATATCATGTGACACTCAT 1137

Qy        277 LysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeu 296
              |||||||||            |||         :::::::::||||||   :::
Db       1138 AAATATGGATTTGCACCAAAAGGCTCGTCAGTTATAATGTATAGAAACAGCGACTTACGA 1197

Qy        297 ProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeu 316
                    :::         :::      :::   |||            :::   ||||||
Db       1198 ATGCATCAGTATTACGTAAATCCTGCTTGGACTGGCGGGTTATATGGCTCTCCTACATTA 1257

Qy        317 AsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGly 336
                    |||:::      :::   |||:::      :::      ::::::   :::|||
Db       1258 GCAGGGTCCAGGCCTGGTGCTATTGTCGTAGGTTGTTGGGCCACTATGGTCAACATGGGT 1317

Qy        337 HisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
                    ||||||         :::|||:::|||:::|||   ::::::
Db       1318 GAAAATGGGTAC---------ATTGAGTCGTGCCAAGAA---ATAGTCGGTGCAGCAATG 1365

Qy        357 GlyLeuGluLysThrGluArgPheAsnIle------------ValSerLysAspGluGly 372
                    :::|||      :::   ||||||            :::      :::
Db       1366 AAGTTTAAAAAATACATCCAGGAAAACATTCCAGACCTGAATATAATGGGCAACCCTAGA 1425

Qy        373 ValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSer 392
                    :::::::::||||||   |||      :::   |||         |||:::|||
Db       1426 TATTCAGTCATTTCATTTTCTTCAAAGACCTTGAACATACAC---------GAACTATCT 1476

Qy        393 AspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGln 412
              |||   |||   :::   ||||||         |||         |||
Db       1477 GACAGGTTGTCCAAGAAAGGCTGGCATTTCAATGCCCTACAAAAGCCG------------ 1524

Qy        413 HisIleThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArg 432
                                |||   :::         |||:::|||   |||:::
Db       1525 ------------------GTTGCACTACACATGGCCTTCACGAGA---TTGAGCGCTCAT 1563

Qy        433 LeuValIleAspIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIle 452
              :::|||   :::|||      ::::::|||
Db       1564 GTTGTGGATGAGATCTGCGACATTTTACGT------------------------------ 1593

Qy        453 HisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSer 466
                       ::::::   :::   ||||||||||||||||||
Db       1594 ------ACTACCGTGCAAGAGTTGAAGAGCGAATCAAATTCT 1629

RESULT 8
US-09-356-643B-1
; Sequence 1, Application US/09356643B
; Patent No. 6569666
```

```
 GENERAL INFORMATION:
  APPLICANT: Saba, Julie D.
  TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
  TITLE OF INVENTION:  POLYNUCLEOTIDES AND MODULATING AGENTS AND
  TITLE OF INVENTION:  METHODS OF USE THEREFOR
  FILE REFERENCE: 200116.402C1
  CURRENT APPLICATION NUMBER: US/09/356,643B
  CURRENT FILING DATE:  1999-07-19
  NUMBER OF SEQ ID NOS: 14
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 1
   LENGTH: 1770
   TYPE: DNA
   ORGANISM: S. cerevisiae
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (1)...(1770)
S-09-356-643B-1

lignment Scores:
red. No.:              6.54e-30       Length:        1770
core:                  325.00         Matches:       122
ercent Similarity:     44.53%         Conservative:  98
est Local Similarity:  24.70%         Mismatches:    188
uery Match:            12.43%         Indels:        86
B:                     4              Gaps:          20

S-10-006-852-2 (1-502) x US-09-356-643B-1 (1-1770)

y          3 LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22
             |||         |||   :::            |||      :::
b        316 TTGAGGGGTACCGTAGAAAAGGAAGTCACAAAGGTCAAACAATCGATCGAAGACGAACTA 375

y         23 ValArg------ThrSerLeuProArgPhe---LysMetProGluAsnSerIleProLys 39
             :::|||      :::   |||      |||   ::::::|||   |||   ||||||:::
b        376 ATTAGATCGGACTCTCAGTTAATGAATTTCCCACAGTTGCCATCCAATGGGATACCTCAG 435

y         40 GluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
             :::         :::   :::|||                           :::||||||
b        436 GATGATGTTATTGAAGAGCTAAAT---------------------------AAATTGAAC 468

y         60 LeuAlaSerPheValThrThrTrpMetGluProGluCys--------------------- 72
                            |||   |||   |||   :::
b        469 GACTTGATACCACATACCCAATGGAAGGAAGGAAAGGTCTCTGGTGCCGTTTACCACGGT 528

y         73 ------AspLysLeuIleMetSerSerIleAsn------------------LysAsnTyr 84
                   |||   :::   :::   :::|||                        |||
b        529 GGTGATGATTTGATCCACTTACAAACAATCGCATACGAAAAATATTGCGTTGCCAATCAA 588

y         85 ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle 104
             :::      |||   :::|||         ::::::::::::      |||:::|||:::
b        589 TTACATCCCGATGTCTTTCCTGCCGTACGTAAAATGGAATCCGAAGTGGTTTCTATGGTT 648

y        105 AlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyValGlyThrVal 124
                   :::||||||||||||         ::::::|||      |||      |||
b        649 TTAAGAATGTTTAATGCCCCT---------TCTGATACAGGTTGTGGTACCACAACTTCA 699

y        125 GlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLys 144
             |||   :::|||:::::::::||||||   |||:::   |||            :::
b        700 GGTGGTACAGAATCCTTGCTTTTAGCATGTCTGAGCGCTAAAATGTATGCCCTTCATCAT 759

y        145 ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln 164
             |||                  :::   :::|||   |||:::
b        760 CGT---------------GGAATCACCGAACCAGAAATAATTGCTCCCGTAACTGCACAT 804

y        165 ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer 184
                   ::::::|||   |||   ||||||   ::::::|||:::   |||:::|||
b        805 GCTGGGTTTGACAAAGCTGCTTATTACTTTGGCATGAAGCTACGCCACGTGGAGCTAGAT 864

y        185 GluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle 204
                      |||   :::|||      :::            :::::::::|||||||||
b        865 CCAACGACATATCAAGTGGACCTGGGAAAAGTGAAAAAATTCATCAATAAGAACACAATT 924

Qy       205 CysValAlaAspIleLeuGlySerThrLeuAsn---------GlyGluPheGluAspVal 221
                         ::::::||||||      |||         |||      :::|||:::
Db       925 TTA---------CTGGTCGGTTCCGCTCCAAACTTTCCTCATGGTATTGCCGATGATATT 975

Qy       222 LysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle 241
             :::   |||         :::   :::|||   |||                  |||:::
Db       976 GAAGGATTGGGTAAA---ATAGCACAAAAATATAAA---------------CTTCCTTTA 1017

Qy       242 HisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu------------TyrPro 257
             |||||||||:::      |||   ||||||      |||:::            |||
Db      1018 CACGTCGACAGTTGTCTAGGTTCCTTTATTGTTTCATTTATGGAAAAGGCTGGTTACAAA 1077

Qy       258 GluLeuGlu---TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHis 276
                |||         |||||||||:::|||   |||   ||||||:::         |||
Db      1078 AATCTGCCATTACTTGACTTTAGAGTCCCGGGAGTCACCTCAATATCATGTGACACTCAT 1137

Qy       277 LysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeu 296
             |||||||||            |||         :::::::::||||||   :::
Db      1138 AAATATGGATTTGCACCAAAAGGCTCGTCAGTTATAATGTATAGAAACAGCGACTTACGA 1197

Qy       297 ProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeu 316
                   :::         :::      :::   |||            :::   ||||||
Db      1198 ATGCATCAGTATTACGTAAATCCTGCTTGGACTGGCGGGTTATATGGCTCTCCTACATTA 1257

Qy       317 AsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGly 336
                   |||:::      :::   |||:::      :::      ::::::   :::|||
Db      1258 GCAGGGTCCAGGCCTGGTGCTATTGTCGTAGGTTGTTGGGCCACTATGGTCAACATGGGT 1317

Qy       337 HisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
                   ||||||         :::|||:::|||:::|||   ::::::
Db      1318 GAAAATGGGTAC---------ATTGAGTCGTGCCAAGAA---ATAGTCGGTGCAGCAATG 1365

Qy       357 GlyLeuGluLysThrGluArgPheAsnIle------------ValSerLysAspGluGly 372
                   :::|||      :::   ||||||            :::      :::
Db      1366 AAGTTTAAAAAATACATCCAGGAAAACATTCCAGACCTGAATATAATGGGCAACCCTAGA 1425

Qy       373 ValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSer 392
                   :::::::::||||||   |||      :::   |||         |||:::|||
Db      1426 TATTCAGTCATTTCATTTTCTTCAAAGACCTTGAACATACAC---------GAACTATCT 1476

Qy       393 AspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGln 412
             |||   |||   :::   ||||||         |||         |||
Db      1477 GACAGGTTGTCCAAGAAAGGCTGGCATTTCAATGCCCTACAAAAGCCG------------ 1524

Qy       413 HisIleThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArg 432
                               |||   :::         |||:::|||   |||:::
Db      1525 ------------------GTTGCACTACACATGGCCTTCACGAGA---TTGAGCGCTCAT 1563

Qy       433 LeuValIleAspIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIle 452
             :::|||   :::|||      ::::::|||
Db      1564 GTTGTGGATGAGATCTGCGACATTTTACGT------------------------------ 1593

Qy       453 HisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSer 466
                      ::::::   :::   ||||||||||||||||||
Db      1594 ------ACTACCGTGCAAGAGTTGAAGAGCGAATCAAATTCT 1629

RESULT 9
US-08-939-309-5
; Sequence 5, Application US/08939309
; Patent No. 6423527
;  GENERAL INFORMATION:
;    APPLICANT:  Saba, Julie D.
;    APPLICANT:  Zhou, Jianhui
;    TITLE OF INVENTION:  SPHINGOSINE-1-PHOSPHATE LYASE
;    TITLE OF INVENTION:  POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
;    TITLE OF INVENTION:  METHODS OF USE THEREFOR
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
```

```
    ZIP:  98104
  COMPUTER READABLE FORM:
    MEDIUM TYPE:  Floppy disk
    COMPUTER:  IBM PC compatible
    OPERATING SYSTEM:  PC-DOS/MS-DOS
    SOFTWARE:  PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/939,309
    FILING DATE:  29-SEP-1997
    CLASSIFICATION:  800
  ATTORNEY/AGENT INFORMATION:
    NAME:  David, Maki J.
    REGISTRATION NUMBER:  31,392
    REFERENCE/DOCKET NUMBER:  200116.402
  TELECOMMUNICATION INFORMATION:
    TELEPHONE:  (206) 622-4900
    TELEFAX:  (206) 682-6031
INFORMATION FOR SEQ ID NO:  5:
  SEQUENCE CHARACTERISTICS:
    LENGTH:  1629 base pairs
    TYPE:  nucleic acid
    STRANDEDNESS:  single
    TOPOLOGY:  linear
  FEATURE:
    NAME/KEY:  CDS
    LOCATION:  1..1626
S-08-939-309-5

lignment Scores:
red. No.:                1.34e-22        Length:          1629
core:                    265.50          Matches:         118
ercent Similarity:       41.41%          Conservative:    99
est Local Similarity:    22.52%          Mismatches:      209
uery Match:              10.15%          Indels:          98
B:                       4               Gaps:            22

S-10-006-852-2 (1-502) x US-08-939-309-5 (1-1629)

y           3 LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22
              |||:::            |||   :::   |||:::               |||
b         229 CTAAACGAAGTAAAGGACGAGCTTGAGAAAAGTCTGAGAATTGTGGATCGAAGCACCGAA 288

y          23 ValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAlaAla 42
                    |||::::::|||         :::            :::   :::   ||||||
b         289 TACTTCACTACAATCCCAAGCCATTCAGTTGGAAGAACTGAAGTACTTCGCCTTGCTGCC 348

y          43 ---TyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
                 |||      :::               |||:::|||      |||::::::   |||
b         349 ATCTATGATGATTTGGAAGGACCAGCTTTTTTGGAAGGA------AGAGTATCTGGAGCA 402

y          62 SerPhe---------------------------------------------------Val 64
                 |||
b         403 GTCTTCAATAGAGAAGACGACAAGGACGAACGGGAGATGTATGAGGAGGTGTTCGGAAAA 462

y          65 ThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIleAsnLysAsnTyr 84
                    |||      |||         ||||||
b         463 TTTGCCTGGACCAACCCACTTTGGCCAAAATTG--------------------------- 495

y          85 ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle 104
                             :::|||            ::::::         |||   |||
b         496 ---------------TTCCCTGGAGTGAGAATCATGGAGGCTGAAGTTGTTCGCATGTGT 540

y         105 AlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyValGlyThrVal 124
                 ::::::   |||            ::::::||||||      |||      :::
b         541 TGTAATATGATGAATGGA---------GATTCGGAGACATGT---GGAACTATGTCAACT 588

y         125 GlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLys 144
              |||   |||   :::|||:::||||||   ||||||   :::            |||:::
b         589 GGTGGATCCATTTCAATTCTTTTGGCGTGCCTGGCTCATCGT------------AATCGT 636

y         145 ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln 164
                          |||:::         :::      ::::::      ::::::|||
b         637 CTTTTGAAAAGAGGAGAAAAGTACACAGAG------ATGATTGTCCCATCATCCGTCCAT 690
```

```
Qy        165 ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer 184
                    :::   |||   |||      |||   ::::::::::::::::::   :::
Db        691 GCAGCGTTCTTCAAAGCTGCCGAATGTTTCCGTATCAAAGTTCGCAAGATTCCAGTTGAT 750

Qy        185 GluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle 204
                       :::   :::|||      :::               :::::::::   |||
Db        751 CCTGTTACTTTCAAAGTAGACCTTGTCAAAATGAAAGCCGCAATTAACAAGAGAACA--- 807

Qy        205 CysValAlaAspIleLeuGlySerThrLeuAsn---------GlyGluPheGluAspVal 221
              |||:::      ::::::||||||      |||         |||      :::|||:::
Db        808 TGTATG------TTAGTTGGATCTGCTCCAAACTTTCCATTTGGAACTGTTGATGACATT 861

Qy        222 LysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle 241
              :::   :::      |||   :::|||                  :::|||   |||:::
Db        862 GAAGCTATTGGACAGCTAGGACTTGAA------------------TATGACATCCCAGTT 903

Qy        242 HisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu---TyrProGluLeuGlu 260
              ||||||||||||      |||||||||:::   |||||||||         |||:::
Db        904 CATGTTGATGCTTGTCTTGGTGGTTTCCTTCTTCCATTCCTTGAAGAAGACGAGATTCGC 963

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              :::|||||||||:::|||   |||   ||||||:::         |||||||||||||||
Db        964 TATGACTTCCGTGTTCCTGGTGTATCTTCGATTTCTGCAGATAGTCACAAATACGGACTC 1023

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
                       |||         |||::::::||||||||||||   |||      :::
Db       1024 GCTCCAAAGGGGTCATCAGTTGTTCTTTATCGCAATAAGGAACTTCTTCATAATCAGTAC 1083

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
                          ::::::   |||            :::   |||:::      |||:::
Db       1084 TTCTGTGATGCTGATTGGCAAGGAGGTATCTATGCATCGGCTACTATGGAAGGATCACGC 1143

Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
                          :::         :::      ::::::            |||||||||
Db       1144 GCTGGGCACAACATTGCACTTTGCTGGGCCGCAATGCTTTATCACGCTCAGGAAGGATAC 1203

Qy        341 ------------ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
                          |||   :::::::::      |||:::            :::|||
Db       1204 AAGGCCAATGCTAGAAAGATTGTTGACACTACAAGAAAG------------ATTAGAAAT 1251

Qy        357 GlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuVal 376
              ||||||         :::         :::            :::   |||   :::|||
Db       1252 GGACTTTCAAACATTAAGGGAATCAAATTACAAGGGCCAAGTGAT---GTTTGTATTGTT 1308

Qy        377 AlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArg 396
              :::::::::      |||                  :::         :::   ::::::
Db       1309 AGCTGGACAACCAATGATGGAGTT------GAACTCTACAGATTCCATAACTTCATGAAG 1362

Qy        397 ArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrVal 416
                       |||   :::               |||            ||||||   |||
Db       1363 GAAAAACATTGGCAACTGAATGGACTTCAATTCCCAGCTGGAGTT---CATATCATGGTC 1419

Qy        417 -LeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAs 436
               |||   ::::::                                    |||::::::||
Db       1420 ACTATGAATCATA------------------------------------CTCATCCTGGA 1443

Qy        436 pIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSe 456
              |      |||:::         :::   ||||||                  ::::::||
Db       1444 CTCGCTGAAGCTTTCGTCGCCGATTGCAGAGCTG------------------CAGTTGAG 1485

Qy        456 rLeuGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSe 476
              ||||   :::            :::|||      |||            ||||||:::
Db       1486 TTTGTCAAAA---------GCCACAAACCATCGGAATCCGACAAGACAAGTGAAGCAGCC 1536

Qy        476 rAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLy 496
                     |||                           |||   |||      ||||||
Db       1537 ATCTACGGACTTG---------------------CTCAAAGTATTCCAGACCGATCGCTT 1575

Qy        496 sLysThrSer 499
                  ||||||
Db       1576 GTTCACGAGT 1585
```

```
ESULT 10
S-09-849-180-5
 Sequence 5, Application US/09849180
 Patent No. 6495359
  GENERAL INFORMATION:
       APPLICANT: Saba, Julie D.
                  Zhou, Jianhui
       TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
                          POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
                          METHODS OF USE THEREFOR
       NUMBER OF SEQUENCES: 10
       CORRESPONDENCE ADDRESS:
            ADDRESSEE: Seed Intellectual Property Law Group
            STREET: 701 Fifth Avenue, Suite 6300
            CITY: Seattle
            STATE: Washington
            COUNTRY: USA
            ZIP: 98055
       COMPUTER READABLE FORM:
            MEDIUM TYPE: Floppy disk
            COMPUTER: IBM PC compatible
            OPERATING SYSTEM: PC-DOS/MS-DOS
            SOFTWARE: PatentIn Release #1.0, Version #1.30
       CURRENT APPLICATION DATA:
            APPLICATION NUMBER: US/09/849,180
            FILING DATE: 04-May-2001
            CLASSIFICATION: <Unknown>
       ATTORNEY/AGENT INFORMATION:
            NAME: Pepe, Jeffrey C.
            REGISTRATION NUMBER: 46,985
            REFERENCE/DOCKET NUMBER: 200116.402
       TELECOMMUNICATION INFORMATION:
            TELEPHONE: (206) 622-4900
            TELEFAX: (206) 682-6031
   INFORMATION FOR SEQ ID NO: 5:
       SEQUENCE CHARACTERISTICS:
            LENGTH: 1629 base pairs
            TYPE: nucleic acid
            STRANDEDNESS: single
            TOPOLOGY: linear
       FEATURE:
            NAME/KEY:  CDS
            LOCATION:  1..1626
       SEQUENCE DESCRIPTION: SEQ ID NO: 5:
JS-09-849-180-5

Alignment Scores:
Pred. No.:               1.34e-22       Length:         1629
Score:                   265.50         Matches:        118
Percent Similarity:      41.41%         Conservative:   99
Best Local Similarity:   22.52%         Mismatches:     209
Query Match:             10.15%         Indels:         98
DB:                      4              Gaps:           22

JS-10-006-852-2 (1-502) x US-09-849-180-5 (1-1629)

Qy        3 LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22
            |||:::            |||   :::   |||:::               |||
Db      229 CTAAACGAAGTAAAGGACGAGCTTGAGAAAAGTCTGAGAATTGTGGATCGAAGCACCGAA 288

Qy       23 ValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAlaAla 42
                  |||::::::|||         :::            :::   :::   ||||||
Db      289 TACTTCACTACAATCCCAAGCCATTCAGTTGGAAGAACTGAAGTACTTCGCCTTGCTGCC 348

Qy       43 ---TyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
               |||      :::               |||:::|||      |||::::::   |||
Db      349 ATCTATGATGATTTGGAAGGACCAGCTTTTTGGAAGGA------AGAGTATCTGGAGCA 402

Qy       62 SerPhe-----------------------------------------------------Val 64
               |||
Db      403 GTCTTCAATAGAGAAGACGACAAGGACGAACGGGAGATGTATGAGGAGGTGTTCGGAAAA 462

Qy       65 ThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIleAsnLysAsnTyr 84
```

```
                  |||      |||         ||||||
Db      463 TTTGCCTGGACCAACCCACTTTGGCCAAAATTG--------------------------- 495

Qy       85 ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle 104
                           :::|||            ::::::         |||   |||
Db      496 ---------------TTCCCTGGAGTGAGAATCATGGAGGCTGAAGTTGTTCGCATGTGT 540

Qy      105 AlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyValGlyThrVal 124
               ::::::   |||            ::::::||||||      |||      :::
Db      541 TGTAATATGATGAATGGA---------GATTCGGAGACATGT---GGAACTATGTCAACT 588

Qy      125 GlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLys 144
            |||   |||   :::|||:::||||||   ||||||   :::            |||:::
Db      589 GGTGGATCCATTTCAATTCTTTTGGCGTGCCTGGCTCATCGT------------AATCGT 636

Qy      145 ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln 164
                        |||:::         :::      ::::::      ::::::|||
Db      637 CTTTTGAAAAGAGGAGAAAAGTACACAGAG------ATGATTGTCCCATCATCCGTCCAT 690

Qy      165 ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer 184
                  :::   |||   |||      |||   ::::::::::::::::::   :::
Db      691 GCAGCGTTCTTCAAAGCTGCCGAATGTTTCCGTATCAAAGTTCGCAAGATTCCAGTTGAT 750

Qy      185 GluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle 204
                     :::   :::|||      :::            :::::::::   |||
Db      751 CCTGTTACTTTCAAAGTAGACCTTGTCAAAATGAAAGCCGCAATTAACAAGAGAACA--- 807

Qy      205 CysValAlaAspIleLeuGlySerThrLeuAsn---------GlyGluPheGluAspVal 221
            |||:::      ::::::||||||      |||         |||      :::|||:::
Db      808 TGTATG------TTAGTTGGATCTGCTCCAAACTTTCCATTTGGAACTGTTGATGACATT 861

Qy      222 LysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle 241
            :::   :::      |||   :::|||                  :::|||   |||:::
Db      862 GAAGCTATTGGACAGCTAGGACTTGAA------------------TATGACATCCCAGTT 903

Qy      242 HisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu---TyrProGluLeuGlu 260
            ||||||||||||      |||||||||:::   |||||||||         |||:::
Db      904 CATGTTGATGCTTGTCTTGGTGGTTTCCTTCTTCCATTCCTTGAAGAAGACGAGATTCGC 963

Qy      261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
            :::|||||||||:::|||   |||   ||||||:::         |||||||||||||||
Db      964 TATGACTTCCGTGTTCCTGGTGTATCTTCGATTTCTGCAGATAGTCACAAATACGGACTC 1023

Qy      281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
                     |||         |||::::::||||||||||||   |||      :::
Db     1024 GCTCCAAAGGGGTCATCAGTTGTTCTTTATCGCAATAAGGAACTTCTTCATAATCAGTAC 1083

Qy      301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
                        ::::::   |||            :::   |||:::      |||:::
Db     1084 TTCTGTGATGCTGATTGGCAAGGAGGTATCTATGCATCGGCTACTATGGAAGGATCACGC 1143

Qy      321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
                        :::         :::      ::::::            |||||||||
Db     1144 GCTGGGCACAACATTGCACTTTGCTGGGCCGCAATGCTTTATCACGCTCAGGAAGGATAC 1203

Qy      341 ------------ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
                        |||   :::::::::      |||:::            :::|||
Db     1204 AAGGCCAATGCTAGAAAGATTGTTGACACTACAAGAAAG------------ATTAGAAAT 1251

Qy      357 GlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuVal 376
            ||||||         :::         :::            :::   |||   :::|||
Db     1252 GGACTTTCAAACATTAAGGGAATCAAATTACAAGGGCCAAGTGAT---GTTTGTATTGTT 1308

Qy      377 AlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArg 396
            :::::::::      |||                  :::         :::   ::::::
Db     1309 AGCTGGACAACCAATGATGGAGTT------GAACTCTACAGATTCCATAACTTCATGAAG 1362

Qy      397 ArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrVal 416
                     |||   :::               |||            ||||||   |||
Db     1363 GAAAAACATTGGCAACTGAATGGACTTCAATTCCCAGCTGGAGTT---CATATCATGGTC 1419

Qy      417 -LeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAs 436
             |||   ::::::                                    |||::::::||
```

```
o        1420 ACTATGAATCATA------------------------------------CTCATCCTGGA 1443

y         436 pIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSe 456
              |      |||:::         :::   ||||||                  ::::::||
b        1444 CTCGCTGAAGCTTTCGTCGCCGATTGCAGAGCTG------------------CAGTTGAG 1485

y         456 rLeuGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSe 476
              ||||   :::            :::|||      |||            ||||||:::
b        1486 TTTGTCAAAA---------GCCACAAACCATCGGAATCCGACAAGACAAGTGAAGCAGCC 1536

y         476 rAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLy 496
                     |||                           |||   |||      ||||||
b        1537 ATCTACGGACTTG---------------------CTCAAAGTATTCCAGACCGATCGCTT 1575

y         496 sLysThrSer 499
                  ||||||
b        1576 GTTCACGAGT 1585

ESULT 11
S-09-356-643B-3
 Sequence 3, Application US/09356643B
 Patent No. 6569666
 GENERAL INFORMATION:
  APPLICANT: Saba, Julie D.
  TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
  TITLE OF INVENTION:  POLYNUCLEOTIDES AND MODULATING AGENTS AND
  TITLE OF INVENTION:  METHODS OF USE THEREFOR
  FILE REFERENCE: 200116.402C1
  CURRENT APPLICATION NUMBER: US/09/356,643B
  CURRENT FILING DATE:  1999-07-19
  NUMBER OF SEQ ID NOS: 14
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 3
   LENGTH: 1629
   TYPE: DNA
   ORGANISM: C. elegans
   FEATURE:
   NAME/KEY: CDS
   LOCATION: (1)...(1629)
S-09-356-643B-3

lignment Scores:
red. No.:              1.34e-22       Length:        1629
core:                  265.50         Matches:       118
ercent Similarity:     41.41%         Conservative:  99
est Local Similarity:  22.52%         Mismatches:    209
uery Match:            10.15%         Indels:        98
B:                     4              Gaps:          22

S-10-006-852-2 (1-502) x US-09-356-643B-3 (1-1629)

y           3 LeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSerArgTyr 22
              |||:::            |||   :::   |||:::               |||
b         229 CTAAACGAAGTAAAGGACGAGCTTGAGAAAAGTCTGAGAATTGTGGATCGAAGCACCGAA 288

y          23 ValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAlaAla 42
                    |||::::::|||         :::            :::   :::   ||||||
b         289 TACTTCACTACAATCCCAAGCCATTCAGTTGGAAGAACTGAAGTACTTCGCCTTGCTGCC 348

y          43 ---TyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
                 |||      :::               |||:::|||      |||::::::   |||
b         349 ATCTATGATGATTTGGAAGGACCAGCTTTTTGGAAGGA------AGAGTATCTGGAGCA 402

y          62 SerPhe---------------------------------------------------Val 64
                 |||
b         403 GTCTTCAATAGAGAAGACGACAAGGACGAACGGGAGATGTATGAGGAGGTGTTCGGAAAA 462

y          65 ThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIleAsnLysAsnTyr 84
                    |||      |||         ||||||
b         463 TTTGCCTGGACCAACCCACTTTGGCCAAAATTG--------------------------- 495

y          85 ValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCysValAsnMetIle 104
                             :::|||            ::::::         |||   |||

Db        496 ---------------TTCCCTGGAGTGAGAATCATGGAGGCTGAAGTTGTTCGCATGTGT 540

Qy        105 AlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGlyValGlyThrVal 124
                 ::::::   |||            ::::::||||||      |||      :::
Db        541 TGTAATATGATGAATGGA---------GATTCGGAGACATGT---GGAACTATGTCAACT 588

Qy        125 GlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLysTrpGlnAsnLys 144
              |||   |||   :::|||:::||||||   ||||||   :::            |||:::
Db        589 GGTGGATCCATTTCAATTCTTTTGGCGTGCCTGGCTCATCGT------------AATCGT 636

Qy        145 ArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGlyAlaAsnValGln 164
                          |||:::         :::      ::::::      ::::::|||
Db        637 CTTTTGAAAAGAGGAGAAAAGTACACAGAG------ATGATTGTCCCATCATCCGTCCAT 690

Qy        165 ValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSer 184
                    :::   |||   |||      |||   ::::::::::::::::::   :::
Db        691 GCAGCGTTCTTCAAAGCTGCCGAATGTTTCCGTATCAAAGTTCGCAAGATTCCAGTTGAT 750

Qy        185 GluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIle 204
                       :::   :::|||      :::            :::::::::   |||
Db        751 CCTGTTACTTTCAAAGTAGACCTTGTCAAAATGAAAGCCGCAATTAACAAGAGAACA--- 807

Qy        205 CysValAlaAspIleLeuGlySerThrLeuAsn---------GlyGluPheGluAspVal 221
              |||:::      ::::::||||||      |||         |||      :::|||:::
Db        808 TGTATG------TTAGTTGGATCTGCTCCAAACTTTCCATTTGGAACTGTTGATGACATT 861

Qy        222 LysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrProIle 241
              :::   :::      |||   :::|||                  :::|||   |||:::
Db        862 GAAGCTATTGGACAGCTAGGACTTGAA------------------TATGACATCCCAGTT 903

Qy        242 HisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu---TyrProGluLeuGlu 260
              ||||||||||||      |||||||||:::   |||||||||         |||:::
Db        904 CATGTTGATGCTTGTCTTGGTGGTTTCCTTCTTCCATTCCTTGAAGAAGACGAGATTCGC 963

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              :::|||||||||:::|||   |||   ||||||:::         |||||||||||||||
Db        964 TATGACTTCCGTGTTCCTGGTGTATCTTCGATTTCTGCAGATAGTCACAAATACGGACTC 1023

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
                       |||         |||::::::||||||||||||   |||      :::
Db       1024 GCTCCAAAGGGGTCATCAGTTGTTCTTTATCGCAATAAGGAACTTCTTCATAATCAGTAC 1083

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
                          ::::::   |||            :::   |||:::      |||:::
Db       1084 TTCTGTGATGCTGATTGGCAAGGAGGTATCTATGCATCGGCTACTATGGAAGGATCACGC 1143

Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
                          :::         :::      ::::::            |||||||||
Db       1144 GCTGGGCACAACATTGCACTTTGCTGGGCCGCAATGCTTTATCACGCTCAGGAAGGATAC 1203

Qy        341 ------------ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGlu 356
                          |||   :::::::::      |||:::            :::|||
Db       1204 AAGGCCAATGCTAGAAAGATTGTTGACACTACAAGAAAG------------ATTAGAAAT 1251

Qy        357 GlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuVal 376
              ||||||         :::         :::            :::   |||   :::|||
Db       1252 GGACTTTCAAACATTAAGGGAATCAAATTACAAGGCCAAGTGAT---GTTTGTATTGTT 1308

Qy        377 AlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArg 396
              :::::::::      |||                  :::         :::   ::::::
Db       1309 AGCTGGACAACCAATGATGGAGTT------GAACTCTACAGATTCCATAACTTCATGAAG 1362

Qy        397 ArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrVal 416
                       |||   :::               |||            ||||||   |||
Db       1363 GAAAAACATTGGCAACTGAATGGACTTCAATTCCCAGCTGGAGTT---CATATCATGGTC 1419

Qy        417 -LeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAs 436
               |||   ::::::                                    |||::::::||
Db       1420 ACTATGAATCATA------------------------------------CTCATCCTGGA 1443

Qy        436 pIleGluLysValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSe 456
              |      |||:::         :::   ||||||                  ::::::||
Db       1444 CTCGCTGAAGCTTTCGTCGCCGATTGCAGAGCTG------------------CAGTTGAG 1485
```

```
y         456 rLeuGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSe 476
              ||||   :::            :::|||      |||            ||||||:::
b        1486 TTTGTCAAAA---------GCCACAAACCATCGGAATCCGACAAGACAAGTGAAGCAGCC 1536

y         476 rAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLy 496
                     |||                           |||   |||      ||||||
b        1537 ATCTACGGACTTG---------------------CTCAAAGTATTCCAGACCGATCGCTT 1575

y         496 sLysThrSer 499
                  ||||||
b        1576 GTTCACGAGT 1585

ESULT 12
S-09-056-556-192/c
 Sequence 192, Application US/09056556
 Patent No. 6350456
  GENERAL INFORMATION:
    APPLICANT:  Reed, Steven G.
    APPLICANT:  Skeiky, Yasir A.W.
    APPLICANT:  Dillon, Davin C.
    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE PREVENTION AND              TREATM
    NUMBER OF SEQUENCES:  241
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SEED and BERRY LLP
      STREET:  6300 Columbia Center, 701 Fifth Avenue
      CITY:  Seattle
      STATE:  Washington
      COUNTRY:  USA
      ZIP:  98104-7092
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/09/056,556
      FILING DATE:  07-APR-1998
      CLASSIFICATION:
    ATTORNEY/AGENT INFORMATION:
      NAME:  Maki, David J.
      REGISTRATION NUMBER:  31,392
      REFERENCE/DOCKET NUMBER:  210121.457
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (206) 622-4900
      TELEFAX:  (206) 682-6031
  INFORMATION FOR SEQ ID NO:  192:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  359 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
JS-09-056-556-192

Alignment Scores:
Pred. No.:               1.17e-22       Length:        359
Score:                   257.00         Matches:       44
Percent Similarity:      61.32%         Conservative:  21
Best Local Similarity:   41.51%         Mismatches:    41
Query Match:             9.83%          Indels:        0
DB:                      4              Gaps:          0

JS-10-006-852-2 (1-502) x US-09-056-556-192 (1-359)

Qy        280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299
              |||   |||   |||:::|||:::|||:::||||||      |||   |||||||||:::
Db        359 CTGACCTATCCCGGCGTCGGGTTTGTCGTGTGGCGCGGGCCCGAGCACCTGCCGGAGGAT 300

Qy        300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
              |||:::|||   :::||||||||||||   |||   ||||||||||||||||||||||||
Db        299 CTGGTTTTCCGGGTCAACTACCTCGGCGGCGACATGCCGACCTTCACCCTGAACTTCTCC 240

Qy        320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
```

```
              :::      :::||||||:::   |||||||||      :::|||||||||   :::|||
Db        239 CGTCCCGGTAACCAGGTGGTGGGCCAGTACTACAACTTCCTGCGGCTGGGGCGCGACGGC 180

Qy        340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
              |||      ||||||:::                        |||   :::   |||
Db        179 TATACCAAGGTGATGCAGGCGCTGTCGCACACCGCCCGGTGGCTGGGTGACCAGCTGCGC 120

Qy        360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
              :::   :::         ::::::|||            :::|||:::|||:::
Db        119 GAGGTGGATCATTGCGAGGTGATCTCGGATGGTTCGGCGATCCCGGTGGTCAGCCCGGAC 60

Qy        380 LeuLysAspSerSerCys 385
                 :::         |||
Db         59 ACACGCCACGGCCAGTGC 42

RESULT 13
US-09-072-596-187/c
; Sequence 187, Application US/09072596
; Patent No. 6458366
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Steven G.
;    APPLICANT:  Skeiky, Yasir A.W.
;    APPLICANT:  Dillon, Davin C.
;    APPLICANT:  Campos-Neto, Antonia
;    APPLICANT:  Houghton, Raymond
;    APPLICANT:  Vedvick, Thomas S.
;    APPLICANT:  Twardzik, Daniel R.
;    APPLICANT:  Lodes, Michael J.
;    APPLICANT:  Hendrickson, Ronald C.
;    TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR DIAGNOSIS OF    TUBERCULOSIS
;    NUMBER OF SEQUENCES:  350
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/072,596
;      FILING DATE:  05-MAY-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maki, David J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  210121.417C9
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  187:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-072-596-187

Alignment Scores:
Pred. No.:               1.17e-22       Length:        359
Score:                   257.00         Matches:       44
Percent Similarity:      61.32%         Conservative:  21
Best Local Similarity:   41.51%         Mismatches:    41
Query Match:             9.83%          Indels:        0
DB:                      4              Gaps:          0

US-10-006-852-2 (1-502) x US-09-072-596-187 (1-359)

Qy        280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299
```

```
              |||   |||   |||:::|||:::|||:::||||||      |||   |||||||||:::
b         359 CTGACCTATCCCGGCGTCGGGTTTGTCGTGGCGCGGGCCCGAGCACCTGCCGGAGGAT 300

y         300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
              |||:::|||   :::||||||||||||   |||   ||||||||||||||||||||||||
b         299 CTGGTTTTCCGGGTCAACTACCTCGGCGGCGACATGCCGACCTTCACCCTGAACTTCTCC 240

y         320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
              :::      :::||||||:::   |||||||||      :::|||||||||   :::|||
b         239 CGTCCCGGTAACCAGGTGGTGGGCCAGTACTACAACTTCCTGCGGCTGGGGCGCGACGGC 180

y         340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
              |||      ||||||:::                        |||   :::   |||
b         179 TATACCAAGGTGATGCAGGCGCTGTCGCACACCGCCCGGTGGCTGGGTGACCAGCTGCGC 120

y         360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
              :::   :::         ::::::|||            :::|||:::|||:::
b         119 GAGGTGGATCATTGCGAGGTGATCTCGGATGGTTCGGCGATCCCGGTGGTCAGCCCGGAC 60

y         380 LeuLysAspSerSerCys 385
                 :::         |||
b          59 ACACGCCACGGCCAGTGC 42

ESULT 14
S-08-939-309-1
 Sequence 1, Application US/08939309
 Patent No. 6423527
  GENERAL INFORMATION:
    APPLICANT:  Saba, Julie D.
    APPLICANT:  Zhou, Jianhui
    TITLE OF INVENTION:  SPHINGOSINE-1-PHOSPHATE LYASE
    TITLE OF INVENTION:  POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
    TITLE OF INVENTION:  METHODS OF USE THEREFOR
    NUMBER OF SEQUENCES:  10
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SEED and BERRY LLP
      STREET:  6300 Columbia Center, 701 Fifth Avenue
      CITY:  Seattle
      STATE:  Washington
      COUNTRY:  USA
      ZIP:  98104
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/939,309
      FILING DATE:  29-SEP-1997
      CLASSIFICATION:  800
    ATTORNEY/AGENT INFORMATION:
      NAME:  David, Maki J.
      REGISTRATION NUMBER:  31,392
      REFERENCE/DOCKET NUMBER:  200116.402
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (206) 622-4900
      TELEFAX:  (206) 682-6031
   INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1707 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
,     TOPOLOGY:  linear
,    FEATURE:
,      NAME/KEY:  CDS
,      LOCATION:  1..1704
JS-08-939-309-1

Alignment Scores:
Pred. No.:              3.27e-20      Length:         1707
Score:                  246.50        Matches:        100
Percent Similarity:     40.84%        Conservative:   76
Best Local Similarity:  23.20%        Mismatches:     210
```

```
Query Match:            9.43%         Indels:         45
DB:                     4             Gaps:           12

US-10-006-852-2 (1-502) x US-08-939-309-1 (1-1707)

Qy         22 TyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAla 41
              ||||||:::|||                  :::|||         :::
Db        334 TATGTGAAAACT------------------CTGCCTGCTCAGGGTATGGGCACAGCTGAG 375

Qy         42 AlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
                    :::   :::   :::         :::||||||:::   :::         |||
Db        376 GTTCTGGAGAGACTCAAGGAGTACAGCTCCATGGATGGTTCCTGGCAAGAAGGGAAAGCC 435

Qy         62 SerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIleAsn 81
              |||                  ||||||:::      :::|||::::::   :::
Db        436 TCAGGAGCTGTGTACAATGGGGAACCGAAGCTCACGGAGCTGCTGGTGCAGGCTTATGGA 495

Qy         82 Lys------------AsnTyrValAspMetAspGluTyrProValThrThrGluLeuGln 97
              :::            |||   :::      |||   :::|||         :::|||:::
Db        496 GAATTCACGTGGAGCAATCCACTGCATCCAGATATCTTCCCTGGATTGCGGAAGTTAGAG 555

Qy         98 AsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThr 117
                       |||   |||         |||||||||
Db        556 GCAGAAATCGTTAGGATGACTTGTTCCCTCTTCAATGGG------------GGACCAGAT 603

Qy        118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137
              :::   |||      |||   |||   :::|||:::|||::::::|||      |||:::
Db        604 TCCTGTGGATGTGTGACTTCTGGGGGAACGGAAAGCATCCTGATGGCCTGCAAAGCTTAC 663

Qy        138 LysArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIle 157
                                :::      |||      |||   :::      |||   |||
Db        664 ------------------CGGGACTTGGCGTTAGAGAAGGGGATCAAAACTCCAGAAATT 705

Qy        158 ValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGlu 177
              |||         :::            ::::::|||   |||   ||||||   ::::::
Db        706 GTGGCTCCCGAGAGTGCCCATGCTGCATTCGACAAAGCAGCTCATTATTTTGGGATGAAG 765

Qy        178 LeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAsp 197
              :::      |||   |||   :::         |||   |||   |||
Db        766 ATTGTCCGAGTTGCACTGAAAAAGAACATGGAGGTG---GATGTGCAGGCAATGAAGAGA 822

Qy        198 MetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGlu 217
                 :::      ||||||      :::                        :::|||
Db        823 GCCATCTCCAGGAACACAGCTATGCTGGTCTGTTCTACCCCACAGTTTCCTCATGGTGTG 882

Qy        218 PheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrp 237
                 :::   |||      :::      |||   |||                     :::
Db        883 ATGGATCCTGTCCCCGAAGTGGCCAAGTTAACTGTCAGA------------------TAT 924

Qy        238 AspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu------ 255
                    |||:::||||||||||||      |||||||||:::      |||:::
Db        925 AAAATCCCACTCCATGTGGATGCTTGTCTGGGGGGCTTCCTCATTGTCTTCATGGAGAAA 984

Qy        256 ------TyrPro---GluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsn 272
                    ||||||   |||      :::|||||||||:::      |||   ||||||:::
Db        985 GCAGGGTACCCACTGGAGAAACCATTTGATTTCCGGGTGAAAGGTGTGACCAGCATTTCA 1044

Qy        273 ValSerGlyHisLysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsn 292
                       ||||||||||||            |||         |||::::::   |||
Db       1045 GCAGATACTCATAAGTATGGCTATGCTCCTAAAGGTTCATCAGTGGTGATGTACTCTAAC 1104

Qy        293 LysGluAspLeuProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnPro 312
              ::::::            :::               ::::::   |||
Db       1105 GAGAAGTACAGGACGTACCAGTTCTTTGTTGGTGCAGACTGGCAAGGTGGTGTCTACGCA 1164

Qy        313 ThrPheThrLeuAsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeu 332
              :::   ::::::      |||:::            :::   |||   :::      |||
Db       1165 TCTCCAAGCATAGCTGGCTCACGGCCTGGTGGCATCATTGCAGCCTGTTGGGCGGCCTTG 1224

Qy        333 IleArgLeuGlyHisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIle 352
              :::      |||      ||||||            :::         :::
Db       1225 ATGCACTTCGGTGAGAACGGCTATGTTGAAGCTACCAAACAGATCATCAAAACTGCTCGC 1284
```

```
y            353 ValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGly 372
                    |||:::      ||||||      :::         |||         |||
b           1285 TTCCTGAAGTCAGAACTGGAAAACATCAAAAACATCTTCATTTTCGGT---GATCCTCAA 1341

y            373 ValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSer 392
                 :::   ::::::|||            |||                  :::   :::|||
b           1342 TTGTCAGTTATTGCTCTGGGATCCAACGATTTTGACATT---------TACCGACTATCT 1392

y            393 AspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGln 412
                 :::|||:::         ||||||            |||         |||   :::
b           1393 AATATGATGTCTGCTAAGGGGTGGAATTTT---AACTACCTGCAGTTCCCAAGAAGCATT 1449

y            413 HisIleThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArg 432
                 |||      :::      |||      |||                        :::|||
b           1450 CATTTCTGCATTACGTTAGTACATACTCGG------------------------AAGCGA 1485

y            433 LeuValIleAspIleGluLysValMetArgGlu 443
                 :::   |||         |||   :::||||||
b           1486 GTGGCGATCCAGTTCCTAAAGGATATCCGGGAA 1518

ESULT 15
S-09-849-180-1
 Sequence 1, Application US/09849180
 Patent No. 6495359
  GENERAL INFORMATION:
        APPLICANT: Saba, Julie D.
                   Zhou, Jianhui
        TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
                            POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
                            METHODS OF USE THEREFOR
        NUMBER OF SEQUENCES: 10
        CORRESPONDENCE ADDRESS:
             ADDRESSEE: Seed Intellectual Property Law Group
             STREET: 701 Fifth Avenue, Suite 6300
             CITY: Seattle
             STATE: Washington
             COUNTRY: USA
             ZIP: 98055
        COMPUTER READABLE FORM:
             MEDIUM TYPE: Floppy disk
             COMPUTER: IBM PC compatible
             OPERATING SYSTEM: PC-DOS/MS-DOS
             SOFTWARE: PatentIn Release #1.0, Version #1.30
        CURRENT APPLICATION DATA:
             APPLICATION NUMBER: US/09/849,180
             FILING DATE: 04-May-2001
             CLASSIFICATION: <Unknown>
        ATTORNEY/AGENT INFORMATION:
             NAME: Pepe, Jeffrey C.
             REGISTRATION NUMBER: 46,985
             REFERENCE/DOCKET NUMBER: 200116.402
        TELECOMMUNICATION INFORMATION:
             TELEPHONE: (206) 622-4900
             TELEFAX: (206) 682-6031
   INFORMATION FOR SEQ ID NO: 1:
        SEQUENCE CHARACTERISTICS:
             LENGTH: 1707 base pairs
             TYPE: nucleic acid
             STRANDEDNESS: single
             TOPOLOGY: linear
        FEATURE:
             NAME/KEY:  CDS
             LOCATION:  1..1704
        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-849-180-1

Alignment Scores:
Pred. No.:              3.27e-20       Length:         1707
Score:                  246.50         Matches:        100
Percent Similarity:     40.84%         Conservative:   76
Best Local Similarity:  23.20%         Mismatches:     210
Query Match:            9.43%          Indels:         45
DB:                     4              Gaps:           12

US-10-006-852-2 (1-502) x US-09-849-180-1 (1-1707)

Qy            22 TyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGluAla 41
                 ||||||:::|||                  :::|||         :::
Db           334 TATGTGAAAACT------------------CTGCCTGCTCAGGGTATGGGCACAGCTGAG 375

Qy            42 AlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAla 61
                       :::   :::   :::         :::||||||:::   :::         |||
Db           376 GTTCTGGAGAGACTCAAGGAGTACAGCTCCATGGATGGTTCCTGGCAAGAAGGGAAAGCC 435

Qy            62 SerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIleAsn 81
                 |||                  ||||||:::      :::|||::::::   :::
Db           436 TCAGGAGCTGTGTACAATGGGGAACCGAAGCTCACGGAGCTGCTGGTGCAGGCTTATGGA 495

Qy            82 Lys------------AsnTyrValAspMetAspGluTyrProValThrThrGluLeuGln 97
                 :::            |||   :::      |||   :::|||         :::|||:::
Db           496 GAATTCACGTGGAGCAATCCACTGCATCCAGATATCTTCCCTGGATTGCGGAAGTTAGAG 555

Qy            98 AsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThr 117
                          |||   |||         |||||||||
Db           556 GCAGAAATCGTTAGGATGACTTGTTCCCTCTTCAATGGG------------GGACCAGAT 603

Qy           118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137
                 :::   |||      |||   |||   :::|||:::|||::::::|||      |||:::
Db           604 TCCTGTGGATGTGTGACTTCTGGGGAACGGAAAGCATCCTGATGGCCTGCAAAGCTTAC 663

Qy           138 LysArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIle 157
                                   :::      |||      |||   :::      |||   |||
Db           664 ------------------CGGGACTTGGCGTTAGAGAAGGGGATCAAAACTCCAGAAATT 705

Qy           158 ValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGlu 177
                 |||         :::            ::::::|||   |||   ||||||   ::::::
Db           706 GTGGCTCCCGAGAGTGCCCATGCTGCATTCGACAAAGCAGCTCATTATTTGGGATGAAG 765

Qy           178 LeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAsp 197
                 :::      |||   |||   :::         |||   |||   |||
Db           766 ATTGTCCGAGTTGCACTGAAAAAGAACATGGAGGTG---GATGTGCAGGCAATGAAGAGA 822

Qy           198 MetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGlu 217
                    :::      ||||||      :::                        :::|||
Db           823 GCCATCTCCAGGAACACAGCTATGCTGGTCTGTTCTACCCCACAGTTTCCTCATGGTGTG 882

Qy           218 PheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrp 237
                    :::   |||      :::      |||   |||                     :::
Db           883 ATGGATCCTGTCCCCGAAGTGGCCAAGTTAACTGTCAGA------------------TAT 924

Qy           238 AspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeu------ 255
                       |||:::||||||||||||      |||||||||:::      |||:::
Db           925 AAAATCCCACTCCATGTGGATGCTTGTCTGGGGGGCTTCCTCATTGTCTTCATGGAGAAA 984

Qy           256 ------TyrPro---GluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsn 272
                       ||||||   |||      :::|||||||||:::      |||   ||||||:::
Db           985 GCAGGGTACCCACTGGAGAAACCATTTGATTTCCGGGTGAAAGGTGTGACCAGCATTTCA 1044

Qy           273 ValSerGlyHisLysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsn 292
                          ||||||||||||            |||         |||::::::   |||
Db          1045 GCAGATACTCATAAGTATGGCTATGCTCCTAAAGGTTCATCAGTGGTGATGTACTCTAAC 1104

Qy           293 LysGluAspLeuProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnPro 312
                 ::::::            :::               ::::::   |||
Db          1105 GAGAAGTACAGGACGTACCAGTTCTTTGTTGGTGCAGACTGGCAAGGTGGTGTCTACGCA 1164

Qy           313 ThrPheThrLeuAsnPheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeu 332
                 :::   ::::::      |||:::            :::   |||   :::      |||
Db          1165 TCTCCAAGCATAGCTGGCTCACGGCCTGGTGGCATCATTGCAGCCTGTTGGGCGGCCTTG 1224

Qy           333 IleArgLeuGlyHisGluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIle 352
                 :::      |||      ||||||            :::         :::
Db          1225 ATGCACTTCGGTGAGAACGGCTATGTTGAAGCTACCAAACAGATCATCAAAACTGCTCGC 1284

Qy           353 ValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAspGluGly 372
                    |||:::      ||||||      :::         |||         |||
```

```
o       1285 TTCCTGAAGTCAGAACTGGAAAACATCAAAAACATCTTCATTTTCGGT---GATCCTCAA 1341

y        373 ValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGluIleSer 392
             :::   ::::::|||            |||                  :::   :::|||
o       1342 TTGTCAGTTATTGCTCTGGGATCCAACGATTTTGACATT---------TACCGACTATCT 1392

y        393 AspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGln 412
             :::|||:::         ||||||            |||         |||   :::
b       1393 AATATGATGTCTGCTAAGGGGTGGAATTTT---AACTACCTGCAGTTCCCAAGAAGCATT 1449

y        413 HisIleThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAlaGluArg 432
             |||      :::      |||      |||                        :::|||
b       1450 CATTTCTGCATTACGTTAGTACATACTCGG------------------------AAGCGA 1485

y        433 LeuValIleAspIleGluLysValMetArgGlu 443
             :::   |||         |||   :::||||||
b       1486 GTGGCGATCCAGTTCCTAAAGGATATCCGGGAA 1518

earch completed: October 22, 2003, 18:12:40
ob time : 2696 secs
```

```
                  GenCore version 5.1.6
        Copyright (c) 1993 - 2003  Compugen Ltd.

M protein - nucleic search, using frame_plus_p2n model

un on:        October 22, 2003, 15:08:39 ; Search time 358 Seconds
                                          (without alignments)
                                          3785.246 Million cell updates/sec

itle:         US-10-006-852-2
erfect score: 2615
equence:      1 MVLSHAVSESDVSVHSTFAS..........DIITGWKKFVADRKKTSGIC 502

coring table: BLOSUM62
              Xgapop 10.0 , Xgapext  0.5
              Ygapop 10.0 , Ygapext  0.5
              Fgapop  6.0 , Fgapext  7.0
              Delop   6.0 , Delext   7.0

earched:      2552756 seqs, 1349719017 residues

otal number of hits satisfying chosen parameters:      5105512

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

Command line parameters:
 MODEL=frame+_p2n.model -DEV=xlp
 Q=/cgn2_1/USPTO_spool_p/US10006852/runat_21102003_164103_2443/app_query.fasta_1.647
 DB=N_Geneseq_19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
 LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
 LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
 MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 USER=US10006852_@CGN_1_1_490_@runat_21102003_164103_2443 -NCPU=6 -ICPU=3
 NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
 DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
 FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :     N_Geneseq_19Jun03:*
              1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
              2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
              3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
              4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
              5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
              6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
              7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
              8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
              9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
              10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
              11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
              12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
              13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
              14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
              15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
              16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
              17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
              18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
              19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
              20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
              21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
              22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
              23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
              24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
              25:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                 SUMMARIES
                     %
 Result            Query
   No.    Score    Match Length DB  ID                 Description
 ---------------------------------------------------------------------------
     1     2615    100.0   1509  24  AAL43410           A thaliana GAD1 co
     2     2607     99.7   1509  24  ABZ14005           Arabidopsis thalia
     3     2240     85.7   1705  24  AAL43415           Tobacco GAD1 codin
     4     2225     85.1   1784  24  AAL43417           Petunia GAD coding
     5     2199     84.1   1771  24  AAL43416           Tobacco GAD2 codin
     6   2135.5     81.7   1482  24  ABZ13132           Arabidopsis thalia
     7     2130     81.5   1665  24  AAL43411           A thaliana GAD2 co
     8     2014     77.0   1783  15  AAQ74691           Early Ripening Tom
     9     2014     77.0   1783  24  AAL43418           Tomato GAD coding
    10     1965     75.1   1605  21  AAC36324           Arabidopsis thalia
    11   1936.5     74.1   2121  24  AAL43413           A thaliana GAD4 co
    12     1836     70.2   1180  22  AAI70004           Tomato glutamate d
    13   1834.5     70.2   2493  24  AAL43412           A thaliana GAD3 co
    14     1782     68.1   1946  24  AAL43414           A thaliana GAD5 co
    15     1570     60.0   1607  21  AAC44889           Arabidopsis thalia
    16     1318     50.4   1545  22  AAH23436           Aspergillus oryzae
    17     1236     47.3   2243  22  AAH23435           Aspergillus oryzae
 c  18     1114     42.6 4403765  22  AAI99683           Mycobacterium tube
 c  19     1114     42.6 4411529  22  AAI99682           Mycobacterium tube
 c  20     1028     39.3 2365589  24  ABA90521           Genomic sequence o
    21     1025     39.2   5565  19  AAV11905           L. lactis NS3 locu
 c  22      971     37.1 684707  24  ABQ67196           Listeria innocua c
 c  23      971     37.1 3011208  24  ABQ69245           Listeria innocua D
    24    959.5     36.7   2742  24  ABQ69218           Listeria monocytog
    25    959.5     36.7   6077  24  ABQ71043           Listeria monocytog
 c  26    958.5     36.7 2944528  24  ABA03041           Listeria monocytog
    27    955.5     36.5   1389  24  ABQ69870           Listeria monocytog
    28    955.5     36.5 2944528  24  ABA03041           Listeria monocytog
    29      950     36.3 684707  24  ABQ67196           Listeria innocua c
    30      950     36.3 3011208  24  ABQ69245           Listeria innocua D
    31      949     36.3   4182  24  ABQ70867           Listeria monocytog
    32      934     35.7    742  19  AAV28662           Ripening banana pu
    33      851     32.5    807  19  AAV28663           Ripening banana pu
    34      697     26.7   3129  23  AAS81862           DNA encoding novel
    35      697     26.7   3129  23  AAS89927           DNA encoding novel
 c  36      697     26.7   3129  23  AAS92819           DNA encoding novel
    37      685     26.2    416  25  ABX62354           Arabidopsis thalia
    38    620.5     23.7    440  24  ABL94157           Arabidopsis thalia
    39      570     21.8    960  24  ABQ75967           Yeast DNA sequence
    40      545     20.8    387  25  ABX19421           Human GDP-mannose
    41    508.5     19.4   1830  23  AAS81857           DNA encoding novel
 c  42      504     19.3   3201  23  AAS93387           DNA encoding novel
    43      504     19.3   3468  23  AAS89926           DNA encoding novel
    44      504     19.3   3944  23  AAS81863           DNA encoding novel
    45      504     19.3   3944  23  AAS88584           DNA encoding novel


                                ALIGNMENTS

RESULT 1
AAL43410
ID   AAL43410 standard; cDNA; 1509 BP.
XX
AC   AAL43410;
XX
DT   25-SEP-2002  (first entry)
XX
DE   A thaliana GAD1 coding sequence.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1509
```

```
T                  /*tag=  a
T                  /product= "GAD1"
X
N    WO200238736-A2.
X
D    16-MAY-2002.
X
F    07-NOV-2001; 2001WO-US47447.
X
R    07-NOV-2000; 2000US-246367P.
X
A    (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I    Kinnersley AM,  Turano FJ;
X
R    WPI; 2002-490073/52.
R    P-PSDB; AA015132.
X
T    Making transformed plants that selectively increase gamma-aminobutyric
T    acid production, by incorporating a DNA construct with a polynucleotide
T    encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T     -
X
S    Claim 18; Page 53; 63pp; English.
X
C    The present invention relates to a method of producing a transformed
C    plant that selectively increases production of gamma-aminobutyric acid
C    (GABA) in response to a signal, by incorporating into the plant's genome
C    a DNA construct with a non-constitutive promoter operably linked to a
C    polynucleotide encoding a functional plant glutamic acid decarboxylase
C    (GAD), to provide a transformed plant that expresses the GAD coding
C    sequence in response to a signal. Plants of this type have an enhanced
C    ability to tolerate environmental or other stresses. The present sequence
C    is the A. thaliana GAD1 coding sequence.
X
Q    Sequence 1509 BP; 433 A; 321 C; 385 G; 370 T; 0 other;

lignment Scores:
red. No.:              2.3e-262       Length:          1509
core:                  2615.00        Matches:         502
ercent Similarity:     100.00%        Conservative:    0
est Local Similarity:  100.00%        Mismatches:      0
uery Match:            100.00%        Indels:          0
B:                     24             Gaps:            0

S-10-006-852-2 (1-502) x AAL43410 (1-1509)

y          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

y         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

y         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

y         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

y         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

y        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

y        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy       141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy       161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy       181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy       201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy       221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy       241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy       261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy       281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy       301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy       321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy       341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy       361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy       381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy       401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy       421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy       441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

Qy       461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

Qy       481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500
```

```
y        501 IleCys 502
             ||||||
b       1501 ATCTGC 1506

ESULT 2
BZ14005
D   ABZ14005 standard; DNA; 1509 BP.
X
C   ABZ14005;
X
T   21-JAN-2003  (first entry)
X
E   Arabidopsis thaliana stress regulated gene SEQ ID NO 1810.
X
W   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
X
S   Arabidopsis thaliana.
X
N   WO200216655-A2.
X
D   28-FEB-2002.
X
F   24-AUG-2001; 2001WO-US26685.
X
R   24-AUG-2000; 2000US-227866P.
R   26-JAN-2001; 2001US-264647P.
R   22-JUN-2001; 2001US-300111P.
X
A   (SCRI ) SCRIPPS RES INST.
A   (SYGN ) SYNGENTA PARTICIPATIONS AG.
X
I   Harper JF,  Kreps J,  Wang X,  Zhu T;
X
R   WPI; 2002-304127/34.
X
T   Identifying a stress condition to which a plant cell has been exposed
T   and producing plants with increased tolerance to these abiotic stresses
T   -
X
S   Claim 144; SEQ ID NO 1810; 577pp + Sequence Listing; English.
X
C   The invention relates to identifying a stress condition to which a plant
C   cell has been exposed, comprising:
C   (a) contacting nucleic acid representative of expressed polynucleotides
C   in the plant cell with an array or probes representative of the plant
C   cell genome; and
C   (b) detecting a profile of expressed polynucleotides in the plant cell
C   characteristic of a stress response. The method is useful in the
C   production of transgenic plants, cells and seeds and in producing plants
C   with increased tolerance to abiotic stress. The present sequence is that
C   of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
C   in methods of the invention.
C   Note: The sequence data for this patent is not represented in the printed
C   specification but is based on sequence information supplied to Derwent by
C   the European Patent Office.
X
Q   Sequence 1509 BP; 432 A; 322 C; 385 G; 370 T; 0 other;

lignment Scores:
red. No.:              1.57e-261     Length:        1509
core:                  2607.00       Matches:       501
ercent Similarity:     99.80%        Conservative:  0
est Local Similarity:  99.80%        Mismatches:    1
uery Match:            99.69%        Indels:        0
B:                     24            Gaps:          0

S-10-006-852-2 (1-502) x ABZ14005 (1-1509)

y          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGAATCCACGGTTGAACTTA 180

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

y          401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

y          421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

y          441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

y          461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

y          481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

y          501 IleCys 502
               ||||||
b         1501 ATCTGC 1506

ESULT 3
AL43415
D   AAL43415 standard; cDNA; 1705 BP.
X
C   AAL43415;
X
T   25-SEP-2002  (first entry)
X
E   Tobacco GAD1 coding sequence.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
X
S   Nicotiana tabacum.
X
H   Key             Location/Qualifiers
T   CDS             71..1561
T                   /*tag=  a
T                   /product= "GAD1"
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I   Kinnersley AM,  Turano FJ;
X
R   WPI; 2002-490073/52.
R   P-PSDB; AAO15137.
X
T   Making transformed plants that selectively increase gamma-aminobutyric
T   acid production, by incorporating a DNA construct with a polynucleotide
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T    -
X
S   Claim 18; Page 56-57; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
C   (GABA) in response to a signal, by incorporating into the plant's genome
C   a DNA construct with a non-constitutive promoter operably linked to a
```

```
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the tobacco GAD1 coding sequence.
XX
SQ   Sequence 1705 BP; 523 A; 301 C; 385 G; 496 T; 0 other;

Alignment Scores:
Pred. No.:              3.19e-223      Length:        1705
Score:                  2240.00        Matches:       428
Percent Similarity:     92.05%         Conservative:  35
Best Local Similarity:  85.09%         Mismatches:    32
Query Match:            85.66%         Indels:        8
DB:                     24             Gaps:          3

US-10-006-852-2 (1-502) x AAL43415 (1-1705)

Qy           1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
               ||||||||||||         ||||||||||||||||||:::||||||||||||||||||
Db          71 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 130

Qy          21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         131 CGATATGTTCGTACTTCTCTTCCGAGGTTTAAGATGCCAGAGAATTCGATACCAAAGGAA 190

Qy          41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         191 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGACTAAATTTA 250

Qy          61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
               ||||||||||||||||||||||||||||||||||||:::||||||:::|||   ||||||
Db         251 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAACAAACTGATGATGGATTCCATT 310

Qy          81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         311 AACAAGAATTACGTTGACATGGATGAATACCCTGTAACCACTGAACTTCAGAATCGATGT 370

Qy         101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
               ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
Db         371 GTAAACATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 430

Qy         121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         431 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 490

Qy         141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
               ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
Db         491 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTGTGACAAGCCCAATATTGTCACTGGT 550

Qy         161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         551 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 610

Qy         181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
               ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
Db         611 GTAAAGTTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 670

Qy         201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         671 GAGAACACAATTTGTGTAGCTGCTATCTTGGGTTCCACACTCAATGGTGAATTTGAAGAT 730

Qy         221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
               ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
Db         731 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 790

Qy         241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         791 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 850

Qy         261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         851 TGGGACTTTAGATTGCCATTGGTGAAGAGTATAAACGTGAGTGGTCACAAATATGGTCTT 910
```

```
y         281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
b         911 GTTTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTACCTGACGAACTT 970

y         301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         971 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1030

y         321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
b        1031 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1090

y         341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              :::||||||||||||||||||:::||||||      ||||||||||||||||||||||||
b        1091 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1150

y         361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
b        1151 AGTGGAAGATTCAACATAATATCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1210

y         381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||:::|||   |||   |||||||||||||||:::   |||||||||:::||||||
b        1211 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1270

y         401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              |||:::|||||||||||||||||||||||||||||||||:::||||||||||||||||||
b        1271 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1330

y         421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        1331 ATTAGAGAAGATTTCTCCCGTACACTCGCCGAGCGACTGGTAATAGACATTGAAAAAGTC 1390

y         441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              :::   |||||||||   ||||||:::||||||      |||:::::::::
b        1391 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTG--------- 1441

y         461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                 :::|||:::|||         :::            ||||||:::|||   :::
b        1442 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1489

y         481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
              |||   :::|||   |||   |||||||||||||||||||||   :::|||||||||:::
b        1490 CAGCTTGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAC 1549

y         500 GlyIleCys 502
              |||:::|||
b        1550 GGAGTTTGT 1558

ESULT 4
AL43417
D   AAL43417 standard; DNA; 1784 BP.
X
C   AAL43417;
X
T   25-SEP-2002  (first entry)
X
E   Petunia GAD coding sequence.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
X
S   Petunia sp.
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
```

```
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   P-PSDB; AAO15139.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
PT   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT    -
XX
PS   Claim 18; Page 58; 63pp; English.
XX
CC   The present invention relates to a method of producing a transformed
CC   plant that selectively increases production of gamma-aminobutyric acid
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the petunia GAD coding sequence.
XX
SQ   Sequence 1784 BP; 550 A; 298 C; 400 G; 536 T; 0 other;

Alignment Scores:
Pred. No.:               1.24e-221      Length:         1784
Score:                   2225.00        Matches:        427
Percent Similarity:      92.45%         Conservative:   38
Best Local Similarity:   84.89%         Mismatches:     34
Query Match:             85.09%         Indels:         5
DB:                      24             Gaps:           2

US-10-006-852-2 (1-502) x AAL43417 (1-1784)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||      ||||||:::||||||||||||:::||||||||||||||||||
Db         72 ATGGTTCTATCAAAGACAGTGTCGCAGAGCGATGTGTCCATTCACTCCACGTTTGCTTCT 131

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db        132 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGATAATTCGATACCAAAAGAA 191

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 GCAGCATATCAGATCATAAATGATGAACTGATGTTAGATGGAAACCCAAGGCTGAACTTG 251

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
Db        252 GCTTCTTTTGTTACAACATGGATGGAACCAGAGTGTGATAAGTTGATGATGGACTCTATT 311

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 AACAAGAACTATGTTGATATGGATGAATATCCTGTTACCACTGAGCTTCAGAATCGATGT 371

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              |||||||||||||||||||||||||||||||||||||||:::   |||||||||||||||
Db        372 GTAAACATGATAGCTCATTTGTTTAATGCACCACTTGAAGATGGAGAAACTGCAGTTGGA 431

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 GTTGGAACTGTTGGATCCTCTGAAGCCATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 491

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
Db        492 TGGCAGAACAAAATGAAAGCCCAAGGCAAACCCTGTGACAAGCCCAACATTGTTACTGGT 551

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 GCAAATGTCCAGGTGTGCTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 611

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||||||||||||||||||||||||||::::::||||||:::|||||||||
```

```
Db        612 GTAAAGCTTAGTGAAGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAGATGGTGGAT 671

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        672 GAAAACACCATTTGTGTAGCTGCTATCTTAGGTTCCACCCTCAATGGAGAATTTGAAGAC 731

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 GTTAAGCGCTTGAATGATCTCTTGGTCGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 791

Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
b         792 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCGTTCATTTACCCAGAGCTTGAG 851

y         261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         852 TGGGACTTTAGATTGCCATTAGTGAAGAGCATTAATGTAAGTGGTCACAAATATGGTCTT 911

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              ||||||||||||||||||||||||:::||||||||||||   |||||||||:::||||||
Db        912 GTCTATGCTGGTATTGGTTGGGTCGTTTGGAGGAACAAGAT-GATTTGCCTGATGAACTT 970

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTTTCTAAA 1030

Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db       1031 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTATGAGGGTTAC 1090

Qy        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              :::||||||||||||||||||:::||||||      ||||||||||||||||||||||||
Db       1091 AAGAATGTGATGGAGAATTGTCAAGAAAATGCATCGGTACTAAGAGAAGGGCTAGAAAAG 1150

Qy        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              |||   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
Db       1151 ACAGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTACCTTTAGTAGCATTCTCTCTT 1210

Qy        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||:::      |||   |||||||||||||||:::   |||||||||:::||||||
Db       1211 AAAGACAACAGGCAACACAACGAGTTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1270

Qy        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 ATTGTTCCTGCATATACTATGCCACCAAACGCACAACACATTACAGTTCTCAGAGTTGTG 1330

Qy        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
b        1331 ATCAGAGAAGATTTCTCCCGTACGCTTGCAGAACGACTGGTAAGAGACATCGAAAAAGTC 1390

Qy        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              :::   |||||||||   ||||||:::||||||      |||:::::::::   :::|||
Db       1391 CTTCATGAACTTGACACACTCCCTGCACGTGTCAATGCTAAGCTCGCTGTGGCCGAGGAG 1450

Qy        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
              ::::::   :::|||         :::            ||||||:::|||   :::
b        1451 CAGGCGGCTGCGAATGGCAGCGAGGTGCAT---------AAGAAAACAGATAGCGAAGTG 1501

Qy        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
              |||   ::::::||||||   |||||||||||||||   :::   :::|||||||||:::
Db       1502 CAGTTGGAGATGATAACTGCATGGAAGAAGTTTGTTGAAGAAAAGAAGAAGAAGACTAAT 1561

Qy        500 GlyIleCys 502
                 :::|||
Db       1562 CGAGTTTGT 1570

RESULT 5
AAL43416
ID   AAL43416 standard; cDNA; 1771 BP.
XX
AC   AAL43416;
XX
DT   25-SEP-2002  (first entry)
```

```
XX
DE   Tobacco GAD2 coding sequence.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
XX
OS   Nicotiana tabacum.
XX
FH   Key             Location/Qualifiers
FT   CDS             67..1557
FT                   /*tag=  a
FT                   /product= "GAD2"
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   P-PSDB; AAO15138.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
PT   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT    -
XX
PS   Claim 18; Page 57; 63pp; English.
XX
CC   The present invention relates to a method of producing a transformed
CC   plant that selectively increases production of gamma-aminobutyric acid
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the tobacco GAD2 coding sequence.
XX
SQ   Sequence 1771 BP; 548 A; 295 C; 393 G; 535 T; 0 other;

Alignment Scores:
Pred. No.:              6.28e-219     Length:         1771
Score:                  2199.00       Matches:        423
Percent Similarity:     90.85%        Conservative:   34
Best Local Similarity:  84.10%        Mismatches:     38
Query Match:            84.09%        Indels:         8
DB:                     24            Gaps:           3

US-10-006-852-2 (1-502) x AAL43416 (1-1771)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||         |||||||||||||||||||||||||||||||||||||||
Db         67 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCGTTCACTCCACTTTCGCCTCC 126

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGAGAATTCAATACCAAAGGAA 186

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GCAGCATATCAGATTATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 246

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              ||||||||||||||||||||||||||||||||||||:::   |||:::|||   ||||||
Db        247 GCATCTTTCGTTACAACATGGATGGAGCCAGAATGTAATACGTTAATGATGGATTCCATT 306
```

```
Qy           81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          307 AACAAGAACTACGTTGACATGGATGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 366

Qy          101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
                ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
b           367 GTAAATATGATAGCTCATTTGTTTAATGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 426

y           121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b           427 GTTGGAACTGTTGGATCCTCTGAAGCTATTATGCTTGCTGGATTAGCCTTTAAGAGAAAA 486

Qy          141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
                ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
b           487 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTTTGATAAGCCCAATATTGTCACCGGT 546

y           161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b           547 GCTAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 606

y           181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
                ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
b           607 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 666

y           201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
                |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b           667 GAGAATACCATTTGTGTTGCTGCTATCTTAGGTTCAACACTCAATGGTGAATTTGAAGAT 726

y           221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
                ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
b           727 GTTAAGCGTTTGAATGACCTTTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 786

y           241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b           787 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 846

y           261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
                |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b           847 TGGGACTTTAGATTGCCATTGGAGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 906

y           281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
                |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
b           907 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 966

y           301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b           967 ATTTTCCACATCAATTACCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1026

y           321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
                ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
b          1027 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1086

y           341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
                :::||||||||||||||||||:::||||||      |||||||||||||||:::||||||
b          1087 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTATTAAGAGAAGGAATTGAAAAA 1146

y           361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
                :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
b          1147 AGTGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTTCCCTTAGTAGCATTTTCTCTT 1206

y           381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
                ||||||:::|||   |||   |||||||||||||||:::   |||||||||:::||||||
Db         1207 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1266

Qy          401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
                ||||||   ||||||||||||||||||||||||||||||:::||||||||||||||||||
Db         1267 ATTGTTCTGGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1326

Qy          421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1327 ATTAGAGAAGATTTCTCCCGCACACTAGCGGAGCGACTGGTAATAGACATTGAAAAAGTC 1386

Qy          441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
```

```
                         :::|||   ||||||:::||||||      |||:::::::::
Db         1387 TTCCACGGAGTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTG--------- 1437

Qy          461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                   :::|||:::|||         :::            ||||||:::|||   :::
Db         1438 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1485

Qy          481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
                |||   :::|||   |||   |||   |||||||||||||||   :::|||||||||:::
Db         1486 CAGCTAGAGATTACTACTGCATGGTTGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAT 1545

Qy          500 GlyIleCys 502
                |||:::|||
Db         1546 GGAGTTTGT 1554

RESULT 6
ABZ13132
ID   ABZ13132 standard; DNA; 1482 BP.
XX
AC   ABZ13132;
XX
DT   21-JAN-2003  (first entry)
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 937.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200216655-A2.
XX
PD   28-FEB-2002.
XX
PF   24-AUG-2001; 2001WO-US26685.
XX
PR   24-AUG-2000; 2000US-227866P.
PR   26-JAN-2001; 2001US-264647P.
PR   22-JUN-2001; 2001US-300111P.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Harper JF,  Kreps J,  Wang X,  Zhu T;
XX
DR   WPI; 2002-304127/34.
XX
PT   Identifying a stress condition to which a plant cell has been exposed
PT   and producing plants with increased tolerance to these abiotic stresses
PT   -
XX
PS   Claim 144; SEQ ID NO 937; 577pp + Sequence Listing; English.
XX
CC   The invention relates to identifying a stress condition to which a plant
CC   cell has been exposed, comprising:
CC   (a) contacting nucleic acid representative of expressed polynucleotides
CC   in the plant cell with an array or probes representative of the plant
CC   cell genome; and
CC   (b) detecting a profile of expressed polynucleotides in the plant cell
CC   characteristic of a stress response. The method is useful in the
CC   production of transgenic plants, cells and seeds and in producing plants
CC   with increased tolerance to abiotic stress. The present sequence is that
CC   of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC   in methods of the invention.
CC   Note: The sequence data for this patent is not represented in the printed
CC   specification but is based on sequence information supplied to Derwent by
CC   the European Patent Office.
XX
SQ   Sequence 1482 BP; 402 A; 331 C; 378 G; 371 T; 0 other;

Alignment Scores:
Pred. No.:             2.02e-212      Length:         1482
Score:                 2135.50        Matches:        411
Percent Similarity:    88.78%         Conservative:   32
```

```
est Local Similarity:  82.36%        Mismatches:   45
uery Match:            81.66%        Indels:       11
B:                     24            Gaps:         2

S-10-006-852-2 (1-502) x ABZ13132 (1-1482)

y           1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||      |||||||||||||||||||||:::||||||||||||||||||
b           1 ATGGTTTTGTCTAAGACAGTTTCCGAATCTGATGTCTCAATCCATTCAACTTTTGCTTCT 60

y          21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||   |||||||||||||||:::|||||||||||||||||||||||||||
b          61 CGTTACGTCCGCAACTCTCTTCCACGATTCGAAATGCCTGAGAACTCAATCCCAAAAGAA 120

y          41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         121 GCAGCTTACCAAATCATCAACGACGAGCTAATGCTCGATGGTAACCCAAGGCTGAACCTA 180

y          61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
b         181 GCTTCCTTCGTGACCACATGGATGGAGCCAGAATGTGACAAGCTCATGATGGAGTCCATC 240

y          81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         241 AACAAGAACTACGTCGACATGGACGAGTACCCTGTCACCACTGAGCTTCAGAACCGATGT 300

y         101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              |||||||||||||||   ||||||||||||||||||   :::   |||   |||||||||
b         301 GTTAACATGATAGCACGTCTCTTCAACGCGCCGCTTGGTGACGGTGAAGCTGCCGTTGGT 360

y         121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
b         361 GTTGGCACCGTCGGATCGTCGGAGGCGATTATGTTGGCCGGTTTGGCTTTTAAGAGACAA 420

y         141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              |||||||||||||||||||||:::|||   |||   ||||||||||||||||||||||||
b         421 TGGCAGAATAAGCGTAAGGCCCAAGGGCTTCCTTATGATAAGCCCAATATCGTAACCGGT 480

y         161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         481 GCTAATGTCCAGGTTTGCTGGGAGAAATTCGCAAGGTATTTCGAAGTGGAGCTTAAGGAA 540

y         181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              |||   |||   |||   ||||||||||||||||||   :::||||||:::|||||||||
b         541 GTGAACCTAAGAGAAGACTATTACGTGATGGACCCTGTAAAGGCGGTCGAAATGGTAGAC 600

y         201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||   |||||||||||||||
b         601 GAAAACACAATTTGTGTCGCTGCCATCCTCGGTTCAACGTTAACCGGTGAATTCGAAGAC 660

y         221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
b         661 GTTAAGCTCCTCAACGACCTCCTTGTCGAGAAAAACAAGCAAACCGGATGGGACACGCCA 720

y         241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         721 ATACACGTGGACGCAGCGAGTGGTGGGTTTATTGCTCCGTTCTTGTATCCGGAGCTGGAG 780

y         261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         781 TGGGATTTCCGGCTACCGTTGGTTAAGAGTATTAATGTGAGTGGTCACAAATACGGTTTG 840

y         281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              ||||||||||||||||||||||||:::||||||   |||   |||||||||:::||||||
b         841 GTTTACGCCGGTATTGGTTGGGTTGTATGGAGAACCAAAACCGATTTGCCTGATGAACTT 900

y         301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         901 ATCTTCCATATCAATTATCTTGGCGCTGATCAACCAACCTTTACACTCAACTTCTCCAAA 960

y         321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
b         961 GGTTCAAGTCAAGTGATTGCTCAGTACTACCAGCTGATTCGTCTTGGATTCGAGGGTTAT 1020

Qy        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              ||||||||||||:::||||||||||||||||||:::|||||||||:::||||||||||||
Db       1021 CGCAATGTGATGGATAATTGTCGGGAAAACATGATGGTACTAAGACAAGGATTAGAGAAA 1080

Qy        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              |||   ||||||   ||||||||||||:::   |||||||||||||||||||||||||||
Db       1081 ACGGGACGTTTTAAAATCGTCTCCAAAGAAAACGGTGTTCCGTTAGTGGCGTTTTCTCTC 1140

Qy        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||||||||   |||   |||||||||::::::      |||||||||:::||||||
Db       1141 AAAGATAGTAGCCGCCACAACGAGTTCGAGGTGGCCCATACACTCCGTCGCTTCGGCTGG 1200

Qy        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              ||||||||||||||||||||||||   :::|||||||||:::||||||||||||||||||
Db       1201 ATCGTTCCGGCCTACACGATGCCTGCGGATGCGCAGCATGTCACTGTCCTTCGAGTTGTT 1260

Qy        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
              ||||||||||||||||||||||||||||||||||||||||||   |||   |||||||||
Db       1261 ATCCGAGAAGATTTCTCTCGAACCTTAGCCGAGAGATTGGTAGCTGATTTCGAGAAGGTT 1320

Qy        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              :::   |||||||||   ||||||:::||||||      |||::::::   |||:::
Db       1321 CTACACGAGCTCGATACGCTTCCGGCGAGGGTTCACGCCAAGATGGCTAATGGAAAAGTT 1380

Qy        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                                                  |||||||||:::      :::
Db       1381 AACGGT------------------------------GTTAAGAAGACGCCAGAGGAGACG 1410

Qy        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSer 499
              ||||||::::::         |||||||||   :::   :::   |||||||||:::
Db       1411 CAGAGAGAAGTCACGGCCTACTGGAAGAAGTTGTTG---GAGACTAAGAAGACCAAC 1464

RESULT 7
AAL43411
ID   AAL43411 standard; cDNA; 1665 BP.
XX
AC   AAL43411;
XX
DT   25-SEP-2002  (first entry)
XX
DE   A thaliana GAD2 coding sequence.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..1501
FT                   /*tag= a
FT                   /product= "GAD2"
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   P-PSDB; AAO15133.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
PT   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT    -
XX
```

```
S   Claim 18; Page 53-54; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
C   (GABA) in response to a signal, by incorporating into the plant's genome
C   a DNA construct with a non-constitutive promoter operably linked to a
C   polynucleotide encoding a functional plant glutamic acid decarboxylase
C   (GAD), to provide a transformed plant that expresses the GAD coding
C   sequence in response to a signal. Plants of this type have an enhanced
C   ability to tolerate environmental or other stresses. The present sequence
C   is the A. thaliana GAD2 coding sequence.
X
Q   Sequence 1665 BP; 490 A; 311 C; 423 G; 441 T; 0 other;

lignment Scores:
red. No.:              8.88e-212      Length:        1665
core:                  2130.00        Matches:       404
ercent Similarity:     89.29%         Conservative:  46
est Local Similarity:  80.16%         Mismatches:    42
uery Match:            81.45%         Indels:        12
B:                     24             Gaps:          3

S-10-006-852-2 (1-502) x AAL43411 (1-1665)

y            1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
               |||||||||:::         :::   :::|||   ||||||   :::   |||   |||
b           17 ATGGTTTTGACAAAAACCGCAACG---AATGATGAATCTGTCTGCACCATGTTCGGATCT 73

y           21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
               |||||||||||||||:::||||||::::::::::::   ||||||||||||||||||:::
b           74 CGCTATGTTCGCACTACACTTCCCAAGTATGAGATTGGTGAGAATTCGATACCGAAAGAC 133

y           41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
               ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
b          134 GCTGCATATCAGATCATAAAAGATGAGCTGATGCTTGATGGTAACCCGAGGCTTAACCTA 193

y           61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
               |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
b          194 GCTTCGTTTGTGACTACATGGATGGAACCAGAGTGTGACAAACTCATCATGGACTCTATC 253

y           81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          254 AACAAGAACTACGTTGATATGGATGAGTACCCTGTCACAACTGAGCTCCAGAACCGATGT 313

y          101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
               ||||||:::||||||   ||||||||||||||||||||||||:::|||||||||||||||
b          314 GTAAACATTATAGCTCGACTGTTCAATGCGCCACTCGAGGAATCTGAGACGGCGGTGGGA 373

y          121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          374 GTAGGGACAGTTGGTTCTTCAGAAGCCATCATGTTAGCCGGATTGGCCTTCAAAAGAAAA 433

y          141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
               |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
b          434 TGGCAGAACAAACGCAAGGCTGAGGGTAAACCCTATGACAAACCCAACATTGTCACTGGA 493

y          161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          494 GCCAATGTTCAAGTTTGCTGGGAGAAATTCGCTCGGTACTTCGAGGTGGAGCTAAAGGAA 553

y          181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
               |||   ||||||||||||||||||||||||||||||   :::|||   :::|||||||||
b          554 GTAAACCTAAGTGAAGGTTACTACGTGATGGATCCAGACAAAGCAGCAGAAATGGTAGAC 613

y          201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b          614 GAGAACACAATCTGTGTCGCAGCCATATTGGGATCCACACTCAACGGTGAGTTCGAAGAC 673

y          221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
               ||||||   ||||||||||||||||||:::||||||:::||||||||||||:::||||||
b          674 GTGAAACGTCTCAATGACTTGCTAGTCAAGAAAAACGAGGAGACTGGTTGGAACACACCG 733

y          241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
               ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db         734 ATCCACGTGGATGCAGCAAGTGGAGGGTTCATAGCTCCGTTTATCTATCCTGAATTAGAA 793

Qy         261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         794 TGGGACTTTAGACTTCCTTTGGTTAAGAGTATCAACGTGAGTGGTCACAAGTATGGACTG 853

Qy         281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
               ||||||||||||||||||||||||:::||||||      |||||||||||||||||||||
Db         854 GTCTATGCTGGTATTGGTTGGGTCGTGTGGAGGGCAGCAGAGGATTTGCCTGAAGAGCTT 913

Qy         301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         914 ATCTTTCATATTAATTATCTTGGTGCTGATCAACCCACTTTCACTCTCAATTTCTCCAAG 973

Qy         321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
               ||||||||||||:::|||||||||||||||||||||||||||||||||   |||||||||
Db         974 GGATCGAGCCAAATTATTGCTCAATACTACCAGCTCATTCGTCTTGGATTCGAGGGGTAC 1033

Qy         341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
               :::||||||||||||||||||   |||||||||:::||||||:::||||||:::||||||
Db        1034 AAAAATGTGATGGAGAATTGCATAGAGAACATGGTGGTTCTCAAAGAAGGGATAGAGAAA 1093

Qy         361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
               ||||||||||||||||||||||||||||||:::|||||||||:::|||||||||||||||
Db        1094 ACAGAGCGTTTCAACATAGTCTCAAAGGACCAAGGAGTGCCAGTCGTAGCCTTCTCTCTC 1153

Qy         381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
               ||||||   |||   |||   |||||||||||||||:::||||||||||||:::||||||
Db        1154 AAGGACCATAGTTTCCACAACGAGTTCGAGATCTCTGAGATGCTACGTCGTTTTGGCTGG 1213

Qy         401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
               ||||||||||||||||||||||||   :::||||||||||||||||||||||||||||||
Db        1214 ATCGTCCCAGCTTACACTATGCCTGCCGATGCACAGCACATCACGGTTCTGCGTGTTGTC 1273

Qy         421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
               ||||||||||||||||||||||||||||||||||||||||||   ||||||   ||||||
Db        1274 ATCAGGGAAGATTTCTCAAGAACACTCGCGGAGAGACTTGTTGCTGATATTTCGAAGGTG 1333

Qy         441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
               :::   |||||||||   |||||||||::::::      |||:::   :::
Db        1334 CTTCATGAGCTAGATACCTTGCCTTCCAAGATATCTAAGAAGATGGGAATA--------- 1384

Qy         461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                                 :::   :::         ||||||:::      ::::::|||
Db        1385 ------------------GAAGGGATCGCGGAAAATGTAAAGGAGAAGAAGATGGAGAAG 1426

Qy         481 Gln------ArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThr 498
               :::         ::::::|||   ||||||:::|||||||||   :::|||||||||
Db        1427 GAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAGTTTGTGAAGGAGAGGAAGAAGATG 1486

Qy         499 SerGlyIleCys 502
               :::|||:::|||
Db        1487 AATGGTGTGTGC 1498

RESULT 8
AAQ74691
ID   AAQ74691 standard; cDNA; 1783 BP.
XX
AC   AAQ74691;
XX
DT   25-MAR-2003  (updated)
DT   22-JUN-1995  (first entry)
XX
DE   Early Ripening Tomato protein ERTD1 gene.
XX
KW   Early Ripening Tomato; transgenic plant; crop improvement;
KW   fruit ripening; transformation; ss.
XX
OS   Lycopersicon esculentum.
XX
PN   WO9421794-A1.
XX
```

```
D   29-SEP-1994.
X
F   22-MAR-1994;   94WO-GB00581.
X
R   22-MAR-1993;   93GB-0005859.
R   22-MAR-1993;   93GB-0005860.
R   22-MAR-1993;   93GB-0005862.
R   22-MAR-1993;   93GB-0005865.
R   22-MAR-1993;   93GB-0005866.
R   22-MAR-1993;   93GB-0005867.
R   22-MAR-1993;   93GB-0005868.
R   22-MAR-1993;   93GB-0005869.
R   12-JUL-1993;   93GB-0014351.
R   12-OCT-1993;   93GB-0020988.
X
A   (ZENE ) ZENECA LTD.
X
I   Abu-Bakar UK,  Barton SL,  Gallego-Veigas PP,  Gray JE;
I   Grierson D,  Lowe AL,  Picton S,  Whotton LC,  Abubakar UK;
X
R   WPI; 1994-317014/39.
X
T   DNA constructs encoding fruit-ripening related proteins - useful
T   for transformation of plants to modify fruit quality
X
S   Claim 1; Page 54; 74pp; English.
X
C   ERTD1 homologous transcripts are 1.8 kb in size.  The transcript
C   is found throughout early fruit development of both wild-type and
C   rin (ripening inhibitor) fruit.  The transcript is detected
C   throughout the entire ripening period in rin fruit.  The transcript
C   is not detected in leaves or wounded leaves of either rin or wild-
C   type tomato.  The sequence has been deposited as NCIMB 40588.  The DNA
C   may be used to transform plants for production of plants with an altered
C   level of expression.  Improvements in transgenic plants include
C   resistance to damage and pests, longer shelf life, improved
C   flavor/aroma, etc.
C   (Updated on 25-MAR-2003 to correct PN field.)
X
Q   Sequence 1783 BP; 512 A; 339 C; 423 G; 509 T; 0 other;

lignment Scores:
red. No.:               1.19e-199      Length:        1783
core:                   2014.00        Matches:       387
ercent Similarity:      86.44%         Conservative:  53
est Local Similarity:   76.03%         Mismatches:    55
uery Match:             77.02%         Indels:        14
B:                      15             Gaps:          5

S-10-006-852-2 (1-502) x AAQ74691 (1-1783)

y          1 MetValLeu---SerHisAlaValSerGluSerAspValSerValHisSerThrPheAla 19
               |||||||||   :::   ::::::   :::|||:::   |||:::|||   |||||||||
b          6 ATGGTGTTAACAACGACGTCGATAAGAGATTCAGAAGAGAGCTTGCACTGTACATTTGCA 65

y         20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
               ||||||||||||:::      ||||||:::||||||||||||:::   |||:::||||||
b         66 TCAAGATATGTACAGGAACCTTTACCTAAGTTCAAAATGCCTAAAAAATCCATGCCGAAA 125

y         40 GluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
               ||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||||
b        126 GAAGCAGCTTATCAGATTGTAAACGACGAGCTTATGTTGGATGGTAACCCCAGGTTGAAT 185

y         60 LeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSer 79
               |||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
b        186 TTAGCTTCCTTTGTTAGCACATGGATGGAGCCCGAGTGCGATAAGCTCATCATGTCATCC 245

y         80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        246 ATTAATAAAAACTATGTCGACATGGATGAGTATCCTGTCACCACTGAACTTCAAAATAGA 305

y        100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaVal 119
               ||||||||||||:::||||||||||||:::||||||:::   :::   ||||||||||||
Db       306 TGTGTTAACATGTTAGCACATCTTTTCCATGCCCCGGTTGGTGATGATGAGACTGCAGTT 365

Qy       120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       366 GGAGTTGGTACAGTGGGTTCATCAGAGGCAATAATGCTTGCTGGCCTTGCTTTCAAACGC 425

Qy       140 LysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThr 159
               |||||||||:::||||||||||||||||||||||||   |||||||||||||||||||||
Db       426 AAATGGCAATCGAAAAGAAAAGCAGAAGGCAAACCTTTCGATAAGCCTAATATAGTCACT 485

Qy       160 GlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLys 179
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       486 GGAGCTAATGTGCAGGTCTGCTGGGAAAAATTTGCAAGGTATTTTGAGGTTGAGTTGAAG 545

Qy       180 GluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetVal 199
               ||||||||||||   ||||||||||||||||||||||||   :::||||||::::::|||
Db       546 GAGGTGAAACTAAAAGAAGGATACTATGTAATGGACCCTGCCAAAGCAGTAGAGATAGTG 605

Qy       200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
               ||||||||||||||||||||||||   ||||||||||||||||||   ||||||||||||
Db       606 GATGAGAATACAATATGTGTTGCTGCAATCCTTGGTTCTACTCTGACTGGGGAGTTTGAG 665

Qy       220 AspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThr 239
               ||||||||||||||||||:::||||||   :::|||||||||||||||||||||:::|||
Db       666 GATGTGAAGCTCCTAAACGAGCTCCTTACAAAAAAGAACAAGGAAACCGGATGGGAGACA 725

Qy       240 ProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeu 259
               ||||||||||||||||||||||||||||||||||||||||||||||||:::|||:::|||
Db       726 CCGATTCATGTCGATGCTGCGAGTGGAGGATTTATTGCTCCTTTCCTCTGGCCAGATCTT 785

Qy       260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       786 GAATGGGATTTCCGTTTGCCTCTTGTGAAAAGTATAAATGTCAGCGGTCACAAGTATGGC 845

Qy       280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299
               |||||||||||||||:::||||||||||||||||||:::|||||||||||||||:::|||
Db       846 CTTGTATATGCTGGTGTCGGTTGGGTGATATGGCGGAGCAAGGAAGACTTGCCCGATGAA 905

Qy       300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
               |||:::|||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       906 CTCGTCTTTCATATAAACTACCTTGGGTCTGATCAGCCTACTTTTACTCTCAACTTCTCT 965

Qy       320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
               |||||||||   |||:::|||||||||||||||||||||||||||||||||   ||||||
Db       966 AAAGGTTCCTATCAAATAATTGCACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGT 1025

Qy       340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
               |||:::|||||||||:::||||||      |||      ||||||   ||||||:::
Db      1026 TATAAGAACGTCATGAAGAATTGCTTATCAAACGCAAAAGTACTAACAGAGGGAATCACA 1085

Qy       360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
               |||      ||||||:::|||||||||||||||   |||||||||:::||||||||||||
Db      1086 AAAATGGGGCGGTTCGATATTGTCTCTAAGGATGTGGGTGTTCCTGTTGTAGCATTTTCT 1145

Qy       380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGly 399
               |||:::|||||||||   :::|||   ||||||:::|||:::   |||||||||:::|||
Db      1146 CTCAGGGACAGCAGCAAATATACGGTATTTGAAGTATCTGAGCATCTCAGAAGATTTGGA 1205

Qy       400 TrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgVal 419
               ||||||||||||||||||||||||||||||:::|||:::||||||   ||||||||||||
Db      1206 TGGATCGTCCCTGCATACACAATGCCACCGGATGCTGAACACATTGCTGTACTGCGGGTT 1265

Qy       420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
               |||||||||||||||||||||   :::||||||||||||||||||   ||||||||||||
Db      1266 GTCATTAGAGAGGATTTCAGCCACAGCCTAGCTGAGAGACTTGTTTCTGACATTGAGAAA 1325

Qy       440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIle------SerLeu 457
               ::::::   |||||||||      |||   |||:::      |||
Db      1326 ATTCTGTCAGAGTTGGACACACAGCCTCCTCGTTTGCCCACCAAAGCTGTCCGTGTCACT 1385

Qy       458 GlyGlnGluLysSerGluSerAsnSerAspAsnLeu------------MetValThrVal 473
                  :::|||      :::         |||   |||            |||   ||||||
Db      1386 GCTGAGGAAGTGCGTGATGACAAGGGTGATGGCTTCATCATTTTCACATGGATACTGTA 1445
```

```
y        474 LysLysSerAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAla 493
                            :::   |||:::|||||||||      |||:::|||
b       1446 ---------------GAGACTCAGAAAGACATTATCAAACATTGGAGGAAAATCGCAGGG 1490

y        494 AspArgLysLysThrSerGlyIleCys 502
                   |||||||||||||||:::|||
b       1491 ------AAGAAGACCAGCGGAGTCTGC 1511

ESULT 9
AL43418
D   AAL43418 standard; cDNA; 1783 BP.
X
C   AAL43418;
X
T   25-SEP-2002  (first entry)
X
E   Tomato GAD coding sequence.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ss.
X
S   Lycopersicon esculentum.
X
H   Key             Location/Qualifiers
T   CDS             6..1514
T                   /*tag=  a
T                   /product= "GAD"
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I   Kinnersley AM,  Turano FJ;
X
R   WPI; 2002-490073/52.
R   P-PSDB; AAO15140.
X
T   Making transformed plants that selectively increase gamma-aminobutyric
T   acid production, by incorporating a DNA construct with a polynucleotide
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T    -
X
S   Claim 18; Page 59; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
C   (GABA) in response to a signal, by incorporating into the plant's genome
C   a DNA construct with a non-constitutive promoter operably linked to a
C   polynucleotide encoding a functional plant glutamic acid decarboxylase
C   (GAD), to provide a transformed plant that expresses the GAD coding
C   sequence in response to a signal. Plants of this type have an enhanced
C   ability to tolerate environmental or other stresses. The present sequence
C   is the tomato GAD coding sequence.
X
Q   Sequence 1783 BP; 512 A; 339 C; 423 G; 509 T; 0 other;

lignment Scores:
red. No.:             1.19e-199     Length:         1783
core:                 2014.00       Matches:        387
ercent Similarity:    86.44%        Conservative:   53
est Local Similarity: 76.03%        Mismatches:     55
uery Match:           77.02%        Indels:         14
B:                    24            Gaps:           5

S-10-006-852-2 (1-502) x AAL43418 (1-1783)
```

```
Qy         1 MetValLeu---SerHisAlaValSerGluSerAspValSerValHisSerThrPheAla 19
             |||||||||   :::   ::::::   :::|||:::   |||:::|||   |||||||||
Db         6 ATGGTGTTAACAACGACGTCGATAAGAGATTCAGAAGAGAGCTTGCACTGTACATTTGCA 65

Qy        20 SerArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLys 39
             ||||||||||||:::      ||||||:::||||||||||||:::   |||:::||||||
Db        66 TCAAGATATGTACAGGAACCTTTACCTAAGTTCAAAATGCCTAAAAAATCCATGCCGAAA 125

Qy        40 GluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsn 59
             ||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||||
Db       126 GAAGCAGCTTATCAGATTGTAAACGACGAGCTTATGTTGGATGGTAACCCCAGGTTGAAT 185

Qy        60 LeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSer 79
             |||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db       186 TTAGCTTCCTTTGTTAGCACATGGATGGAGCCCGAGTGCGATAAGCTCATCATGTCATCC 245

Qy        80 IleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArg 99
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       246 ATTAATAAAAACTATGTCGACATGGATGAGTATCCTGTCACCACTGAACTTCAAAATAGA 305

Qy       100 CysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaVal 119
             ||||||||||||:::||||||||||||:::||||||:::   :::   ||||||||||||
Db       306 TGTGTTAACATGTTAGCACATCTTTTCCATGCCCCGGTTGGTGATGATGAGACTGCAGTT 365

Qy       120 GlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArg 139
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       366 GGAGTTGGTACAGTGGGTTCATCAGAGGCAATAATGCTTGCTGGCCTTGCTTTCAAACGC 425

Qy       140 LysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThr 159
             |||||||||:::||||||||||||||||||||||||   |||||||||||||||||||||
Db       426 AAATGGCAATCGAAAAGAAAAGCAGAAGGCAAACCTTTCGATAAGCCTAATATAGTCACT 485

Qy       160 GlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLys 179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       486 GGAGCTAATGTGCAGGTCTGCTGGGAAAAATTTGCAAGGTATTTTGAGGTTGAGTTGAAG 545

Qy       180 GluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetVal 199
             ||||||||||||   ||||||||||||||||||||||||   :::||||||::::::|||
Db       546 GAGGTGAAACTAAAAGAAGGATACTATGTAATGGACCCTGCCAAAGCAGTAGAGATAGTG 605

Qy       200 AspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGlu 219
             ||||||||||||||||||||||||   ||||||||||||||||||   ||||||||||||
Db       606 GATGAGAATACAATATGTGTTGCTGCAATCCTTGGTTCTACTCTGACTGGGGAGTTTGAG 665

Qy       220 AspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThr 239
             ||||||||||||||||||:::||||||   :::|||||||||||||||||||||:::|||
Db       666 GATGTGAAGCTCCTAAACGAGCTCCTTACAAAAAAGAACAAGGAAACCGGATGGGAGACA 725

Qy       240 ProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeu 259
             ||||||||||||||||||||||||||||||||||||||||||||||||:::|||:::|||
Db       726 CCGATTCATGTCGATGCTGCGAGTGGAGGATTTATTGCTCCTTTCCTCTGGCCAGATCTT 785

Qy       260 GluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGly 279
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       786 GAATGGGATTTCCGTTTGCCTCTTGTGAAAAGTATAAATGTCAGCGGTCACAAGTATGGC 845

Qy       280 LeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGlu 299
             |||||||||||||||:::||||||||||||||||||:::|||||||||||||||:::|||
Db       846 CTTGTATATGCTGGTGTCGGTTGGGTGATATGGCGGAGCAAGGAAGACTTGCCCGATGAA 905

Qy       300 LeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSer 319
             |||:::|||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       906 CTCGTCTTTCATATAAACTACCTTGGGTCTGATCAGCCTACTTTTACTCTCAACTTCTCT 965

Qy       320 LysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGly 339
             |||||||||   |||:::|||||||||||||||||||||||||||||||||   ||||||
Db       966 AAAGGTTCCTATCAAATAATTGCACAGTATTATCAGTTAATAAGACTTGGCTTTGAGGGT 1025

Qy       340 TyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGlu 359
             |||:::|||||||||:::||||||      |||      ||||||   ||||||:::
Db      1026 TATAAGAACGTCATGAAGAATTGCTTATCAAACGCAAAAGTACTAACAGAGGGAATCACA 1085

Qy       360 LysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSer 379
```

```
             |||      ||||||:::|||||||||||||||   |||||||||:::||||||||||||
b       1086 AAAATGGGGCGGTTCGATATTGTCTCTAAGGATGTGGGTGTTCCTGTTGTAGCATTTTCT 1145

y        380 LeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGly 399
             |||:::|||||||||   :::|||   ||||||:::|||:::   |||||||||:::|||
b       1146 CTCAGGGACAGCAGCAAATATACGGTATTTGAAGTATCTGAGCATCTCAGAAGATTTGGA 1205

y        400 TrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgVal 419
             ||||||||||||||||||||||||||||||:::|||:::||||||   ||||||||||||
b       1206 TGGATCGTCCCTGCATACACAATGCCACCGGATGCTGAACACATTGCTGTACTGCGGGTT 1265

y        420 ValIleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLys 439
             |||||||||||||||||||||   :::||||||||||||||||||   ||||||||||||
b       1266 GTCATTAGAGAGGATTTCAGCCACAGCCTAGCTGAGAGACTTGTTTCTGACATTGAGAAA 1325

y        440 ValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIle------SerLeu 457
             ::::::   |||||||||      |||   |||:::      |||
b       1326 ATTCTGTCAGAGTTGGACACACAGCCTCCTCGTTTGCCCACCAAAGCTGTCCGTGTCACT 1385

y        458 GlyGlnGluLysSerGluSerAsnSerAspAsnLeu------------MetValThrVal 473
                :::|||      :::         |||   |||            |||   ||||||
b       1386 GCTGAGGAAGTGCGTGATGACAAGGGTGATGGGCTTCATCATTTTCACATGGATACTGTA 1445

y        474 LysLysSerAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLysPheValAla 493
                            :::   |||:::|||||||||      |||:::|||
b       1446 ---------------GAGACTCAGAAAGACATTATCAAACATTGGAGGAAAATCGCAGGG 1490

y        494 AspArgLysLysThrSerGlyIleCys 502
                   |||||||||||||||:::|||
b       1491 ------AAGAAGACCAGCGGAGTCTGC 1511

ESULT 10
AC36324
D   AAC36324 standard; DNA; 1605 BP.
X
C   AAC36324;
X
T   17-OCT-2000  (first entry)
X
E   Arabidopsis thaliana DNA fragment SEQ ID NO: 13376.
X
W   Hybridisation assay; genetic mapping; gene expression control;
W   protein identification; signal transduction pathway;
W   metabolic pathway; promoter; termination sequence; ss.
X
S   Arabidopsis thaliana.
X
N   EP1033405-A2.
X
D   06-SEP-2000.
X
F   25-FEB-2000; 2000EP-0301439.
X
R   25-FEB-1999;   99US-0121825.
R   05-MAR-1999;   99US-0123180.
R   09-MAR-1999;   99US-0123548.
R   23-MAR-1999;   99US-0125788.
R   25-MAR-1999;   99US-0126264.
R   29-MAR-1999;   99US-0126785.
R   01-APR-1999;   99US-0127462.
R   06-APR-1999;   99US-0128234.
R   08-APR-1999;   99US-0128714.
R   16-APR-1999;   99US-0129845.
R   19-APR-1999;   99US-0130077.
R   21-APR-1999;   99US-0130449.
R   23-APR-1999;   99US-0130510.
R   23-APR-1999;   99US-0130891.
R   28-APR-1999;   99US-0131449.
R   30-APR-1999;   99US-0132048.
R   30-APR-1999;   99US-0132407.
R   04-MAY-1999;   99US-0132484.
R   05-MAY-1999;   99US-0132485.
R   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
```

```
R    22-JUL-1999;    99US-0145192.
R    23-JUL-1999;    99US-0145145.
R    23-JUL-1999;    99US-0145218.
R    23-JUL-1999;    99US-0145224.
R    26-JUL-1999;    99US-0145276.
R    27-JUL-1999;    99US-0145913.
R    27-JUL-1999;    99US-0145918.
R    27-JUL-1999;    99US-0145919.
R    28-JUL-1999;    99US-0145951.
R    02-AUG-1999;    99US-0146386.
R    02-AUG-1999;    99US-0146388.
R    02-AUG-1999;    99US-0146389.
R    03-AUG-1999;    99US-0147038.
R    04-AUG-1999;    99US-0147204.
R    04-AUG-1999;    99US-0147302.
R    05-AUG-1999;    99US-0147192.
R    05-AUG-1999;    99US-0147260.
R    06-AUG-1999;    99US-0147303.
R    06-AUG-1999;    99US-0147416.
R    09-AUG-1999;    99US-0147493.
R    09-AUG-1999;    99US-0147935.
R    10-AUG-1999;    99US-0148171.
R    11-AUG-1999;    99US-0148319.
R    12-AUG-1999;    99US-0148341.
R    13-AUG-1999;    99US-0148565.
R    13-AUG-1999;    99US-0148684.
R    16-AUG-1999;    99US-0149368.
R    17-AUG-1999;    99US-0149175.
R    18-AUG-1999;    99US-0149426.
R    20-AUG-1999;    99US-0149722.
R    20-AUG-1999;    99US-0149723.
R    20-AUG-1999;    99US-0149929.
R    23-AUG-1999;    99US-0149902.
R    23-AUG-1999;    99US-0149930.
R    25-AUG-1999;    99US-0150566.
R    26-AUG-1999;    99US-0150884.
R    27-AUG-1999;    99US-0151065.
R    27-AUG-1999;    99US-0151066.
R    27-AUG-1999;    99US-0151080.
R    30-AUG-1999;    99US-0151303.
R    31-AUG-1999;    99US-0151438.
R    01-SEP-1999;    99US-0151930.
R    07-SEP-1999;    99US-0152363.
R    10-SEP-1999;    99US-0153070.
R    13-SEP-1999;    99US-0153758.
R    15-SEP-1999;    99US-0154018.
R    16-SEP-1999;    99US-0154039.
R    20-SEP-1999;    99US-0154779.
R    22-SEP-1999;    99US-0155139.
R    23-SEP-1999;    99US-0155486.
R    24-SEP-1999;    99US-0155659.
R    28-SEP-1999;    99US-0156458.
R    29-SEP-1999;    99US-0156596.
R    04-OCT-1999;    99US-0157117.
R    05-OCT-1999;    99US-0157753.
R    06-OCT-1999;    99US-0157865.
R    07-OCT-1999;    99US-0158029.
R    08-OCT-1999;    99US-0158232.
R    12-OCT-1999;    99US-0158369.
R    13-OCT-1999;    99US-0159293.
R    13-OCT-1999;    99US-0159294.
R    13-OCT-1999;    99US-0159295.
R    14-OCT-1999;    99US-0159329.
R    14-OCT-1999;    99US-0159330.
R    14-OCT-1999;    99US-0159331.
R    14-OCT-1999;    99US-0159637.
R    14-OCT-1999;    99US-0159638.
R    18-OCT-1999;    99US-0159584.
R    21-OCT-1999;    99US-0160741.
R    21-OCT-1999;    99US-0160767.
R    21-OCT-1999;    99US-0160768.
R    21-OCT-1999;    99US-0160770.
R    21-OCT-1999;    99US-0160814.
```

```
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

Alignment Scores:
Pred. No.:                1.3e-194      Length:         1605
Score:                    1965.00       Matches:        372
Percent Similarity:       86.29%        Conservative:   56
Best Local Similarity:    75.00%        Mismatches:     62
Query Match:              75.14%        Indels:         6
DB:                       21            Gaps:           3

US-10-006-852-2 (1-502) x AAC36324 (1-1605)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              |||||||||:::         |||:::||||||      :::||||||||||||||||||
Db         40 ATGGTACTCGCA---ACCAACTCTGACTCCGACGAGCATTTGCATTCCACTTTTGCTTCT 96

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||      :::||||||||||||||||||::::::   :::||||||:::
Db         97 AGATATGTCCGTGCTGTTGTTCCCAGGTTCAAGATGCCTGACCATTGCATGCCCAAAGAT 156

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db        157 GCTGCTTATCAAGTGATCAATGATGAGTTGATGCTTGATGGTAATCCCAGGCTTAACCTA 216

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||   |||:::
Db        217 GCCTCCTTTGTCACCACTTGGATGGAACCTGAGTGTGACAAACTCATCATGGATTCTGTC 276

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 AATAAGAACTATGTTGATATGGATGAATATCCTGTCACCACTGAGCTCCAGAACCGGTGT 336

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              |||||||||||||||:::   |||:::||||||:::   |||   |||   |||:::|||
Db        337 GTAAATATGATAGCAAACTTTTTCCATGCTCCCGTTGGAGAAGACGAGGCTGCTATTGGG 396

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 TGTGGAACTGTTGGTTCATCTGAGGCTATAATGCTTGCTGGTTTGGCTTTCAAAAGGAAA 456

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||::::::|||||||||:::|||   |||:::||||||||||||||||||||||||
Db        457 TGGCAACATAGGAGAAAAGCTCAGGGTCTACCTATTGATAAGCCTAACATTGTCACTGGA 516

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 GCCAATGTTCAGGTGTGCTGGGAGAAGTTTGCAAGGTACTTTGAGGTAGAGCTCAAAGAG 576

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              |||||||||||||||   ||||||||||||||||||   :::||||||:::|||||||||
Db        577 GTGAAACTAAGTGAAGACTACTATGTTATGGATCCAGCTAAAGCTGTAGAGATGGTGGAT 636

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||   |||||||||||||||
Db        637 GAGAATACCATCTGTGTTGCAGCAATTCTAGGATCTACACTTACTGGAGAGTTTGAGGAC 696

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   |||||||||||||||   |||||||||   ||||||||||||:::||||||
Db        697 GTTAAGCAATTGAACGATCTCTTAGCTGAGAAAAACGCAGAGACAGGATGGGAAACTCCT 756
```

```
y        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
b        757 ATTCATGTTGATGCAGCCAGTGGAGGATTCATTGCTCCTTTCCTCTACCCTGATCTTGAA 816

y        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
b        817 TGGGACTTTAGGCTTCCATGGGTGAAGAGTATTAACGTCAGTGGTCACAAGTATGGACTT 876

y        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             ||||||||||||:::|||||||||:::||||||   |||:::||||||||||||||||||
b        877 GTGTATGCAGGAGTTGGTTGGGTTGTCTGGAGAACAAAAGATGATTTGCCAGAGGAACTT 936

y        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        937 GTCTTCCACATCAACTACTTGGGAGCTGATCAACCCACTTTCACTCTCAACTTCTCAAAA 996

y        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||:::||||||||||||||||||   ||||||||||||   |||||||||
b        997 GGGTCGAGCCAAATCATTGCTCAGTACTATCAGTTTATCCGACTAGGCTTTGAGGGATAC 1056

y        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             :::|||:::||||||||||||   :::|||         ||||||||||||:::|||
b       1057 AAGAACATAATGGAAAACTGCATGGATAACGCAAGGAGGCTAAGAGAAGGAATAGAGATG 1116

y        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
             |||   :::|||||||||||||||||||||   |||||||||||||||||||||||||||
b       1117 ACAGGGAAGTTCAACATTGTGTCCAAAGATATTGGCGTGCCACTAGTGGCATTCTCTCTC 1176

y        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
             ||||||||||||   ||||||   |||||||||::::::   ||||||::::::||||||
b       1177 AAAGACAGTAGCAAGCACACGGTGTTTGAGATCGCAGAGTCTTTGAGAAAATTCGGGTGG 1236

y        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
             |||:::||||||||||||||||||   :::||||||||||||   |||||||||||||||
b       1237 ATCATACCGGCTTACACTATGCCTGCAGATGCACAGCACATTGCTGTGCTCAGAGTTGTG 1296

y        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
             |||||||||||||||||||||   ||||||:::||||||:::      |||   :::|||
b       1297 ATAAGAGAAGACTTTAGCCGAGGCCTTGCAGATAGACTCATCACACATATCATTCAGGTG 1356

y        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
             ::::::|||::::::   ||||||||||||:::   |||      :::
b       1357 CTGAAAGAGATTGAAGGGCTTCCTAGCAGGATTGCACATCTTGCTGCG------------ 1404

y        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                :::      :::   |||:::      |||   ||||||   :::   :::
b       1405 GCTGCAGCGGTTAGTGGTGATGATGAAGAAGTTAAAGTGAAGACTGCCAAGATGTCCTTG 1464

y        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLys 496
             :::   ||||||         ||||||:::   |||      ::::::
b       1465 GAG---GATATCACTAAGTATTGGAAACGCCTTGTGGAACACAAGAGA 1509

ESULT 11
AL43413
D   AAL43413 standard; DNA; 2121 BP.
X
C   AAL43413;
X
T   25-SEP-2002  (first entry)
X
E   A thaliana GAD4 coding sequence.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
X
S   Arabidopsis thaliana.
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
```

```
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   P-PSDB; AAO15135.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
PT   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT    -
XX
PS   Claim 18; Page 55; 63pp; English.
XX
CC   The present invention relates to a method of producing a transformed
CC   plant that selectively increases production of gamma-aminobutyric acid
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the A. thaliana GAD4 coding sequence.
XX
SQ   Sequence 2121 BP; 626 A; 412 C; 473 G; 610 T; 0 other;

Alignment Scores:
Pred. No.:              1.8e-191      Length:         2121
Score:                  1936.50       Matches:        410
Percent Similarity:     61.99%        Conservative:   32
Best Local Similarity:  57.50%        Mismatches:     46
Query Match:            74.05%        Indels:         226
DB:                     24            Gaps:           7

US-10-006-852-2 (1-502) x AAL43413 (1-2121)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||      |||||||||||||||||||||:::||||||||||||||||||
Db          1 ATGGTTTTGTCTAAGACAGTTTCCGAATCTGATGTCTCAATCCATTCAACTTTTGCTTCT 60

Qy         21 ArgTyrValArgThrSerLeuPro------------------------------------ 28
              ||||||||||||   |||||||||
Db         61 CGTTACGTCCGCAACTCTCTTCCACGGTAACAACTTGTAACACAAATCTTTTGTCTAATG 120

Qy         28 ------------------------------------------------------------ 28

Db        121 TTTTCGTCAACAATAGTAACATGTAATGATGTAAACCTTGGATAGTTTTTTTTTGGCCG 180

Qy         28 ------------------------------------------------------------ 28

Db        181 TGGTTAATGTTGTAGATTTATTATGTGTTATATACTATAAGGAAGGACATGTTTCGTTAT 240

Qy         29 ------------------------------ArgPheLysMetProGluAsnSerIlePro 38
                                            ||||||:::|||||||||||||||||||||
Db        241 TTTAACTTAATGTATCATCATTTCATCATTAGATTCGAAATGCCTGAGAACTCAATCCCA 300

Qy         39 LysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeu 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGAAGCAGCTTACCAAATCATCAACGACGAGCTAATGCTCGATGGTAACCCAAGGCTG 360

Qy         59 AsnLeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSer 78
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db        361 AACCTAGCTTCCTTCGTGACCACATGGATGGAGCCAGAATGTGACAAGCTCATGATGGAG 420

Qy         79 SerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsn 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TCCATCAACAAGAACTACGTCGACATGGACGAGTACCCTGTCACCACTGAGCTTCAGAAC 480

Qy         99 ArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAla 118
              |||||||||||||||||||||   ||||||||||||||||||   :::   |||   |||
Db        481 CGATGTGTTAACATGATAGCACGTCTCTTCAACGCGCCGCTTGGTGACGGTGAAGCTGCC 540
```

```
     119 ValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLys 138
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     541 GTTGGTGTTGGCACCGTCGGATCGTCGGAGGCGATTATGTTGGCCGGTTTGGCTTTTAAG 600

     139 ArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleVal 158
         |||:::|||||||||||||||||||||:::|||   |||   ||||||||||||||||||
     601 AGACAATGGCAGAATAAGCGTAAGGCCCAAGGGCTTCCTTATGATAAGCCCAATATCGTA 660

     159 ThrGlyAlaAsnVal--------------------------------------------- 163
         |||||||||||||||
     661 ACCGGTGCTAATGTCCAGGTAAACCAAAACAAAAATTGATGAAATATTAACCAAGACAAA 720

     164 -------------------------------------------------------GlnVa 165
                                                                |||||
     721 ATTGAATTTATCAATCCGGTTAAGTTATATGTGTGACTCAATTTCCGGTTCAATACAGGT 780

     165 lCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGluValLysLeuSerGl 185
         |||||||||||||||||||||||||||||||||||||||||||||||||   |||   ||
     781 TTGCTGGGAGAAATTCGCAAGGTATTTCGAAGTGGAGCTTAAGGAAGTGAACCTAAGAGA 840

     185 uGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAspGluAsnThrIleCy 205
         |   ||||||||||||||||||   :::||||||:::|||||||||||||||||||||||
     841 AGACTATTACGTGATGGACCCTGTAAAGGCGGTCGAAATGGTAGACGAAAACACAATTTG 900

     205 sValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAspValLysLeuLeuAs 225
         |||||||   ||||||||||||||||||   |||||||||||||||||||||||||||||
     901 TGTCGCTGCCATCCTCGGTTCAACGTTAACCGGTGAATTCGAAGACGTTAAGCTCCTCAA 960

     225 nAspLeuLeuValGluLysAsnLysGluThrGly-------------------------- 236
         |||||||||||||||||||||||||:::|||
     961 CGACCTCCTTGTCGAGAAAAACAAGCAAACCGG-GTAATTAAACCAAACCGAGAAACAAG 1019

     236 ------------------------------------------------------------ 236

    1020 CTAATATCGATTGTAATCGGTTTGGAGTCCGGTTTTAACGTTCTAAAACACAATTTGCAG 1079

     237 -TrpAspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTy 256
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    1080 ATGGGACACGCCAATACACGTGGACGCAGCGAGTGGTGGGTTTATTGCTCCGTTCTTGTA 1139

     256 rProGluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHi 276
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    1140 TCCGGAGCTGGAGTGGGATTTCCGGCTACCGTTGGTTAAGAGTATTAATGTGAGTGGTCA 1199

     276 sLysTyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLe 296
         |||||||||||||||||||||||||||||||||||||:::||||||   |||   |||||
    1200 CAAATACGGTTTGGTTACGCCGGTATTGGTTGGGTTGTATGGAGAACCAAAACCGATTT 1259

y    296 uProGluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLe 316
         ||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||
b   1260 GCCTGATGAACTTATCTTCCATATCAATTATCTTGGCGCTGATCAACCAACCTTTACACT 1319

y    316 uAsnPheSerLys----------------------------------------------- 320
         |||||||||||||
b   1320 CAACTTCTCCAAAGGTACATTACCATAAGTCCATAACATATATAACTTTCAATAATATTT 1379

y    321 -------------------------------------------------------GlySe 322
                                                                |||||
b   1380 TTGGTGTATGGAATTGTTTTATAGACTAAACATTTGATAATGCTTGTATAAACCAGGTTC 1439

y    322 rSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGlu----------- 338
         |||||||||||||||||||||||||||||||||||||||||||   |||
b   1440 AAGTCAAGTGATTGCTCAGTACTACCAGCTGATTCGTCTTGGATTCGAGGTAAATAATAA 1499

y    338 ------------------------------------------------------------ 338

b   1500 CTCAAAATAGCAATATATTTACCAAATGGTCAATAAAGAAACTAGAATGTATTATATTTA 1559

y    338 ------------------------------------------------------------ 338

b   1560 AGTTGTTACTTGTTACTATACTTTGAATTAAACGTTCCTAACATGACTAGTTTTGGTATT 1619
```

```
Qy       339 ----------------------------------GlyTyrArgAsnValMetGluAsnCy 347
                                               ||||||||||||||||||:::|||||
Db      1620 GTGTAATTAATAATGTTTTTCTTGTTTGATTTAGGGTTATCGCAATGTGATGGATAATTG 1679

Qy       347 sArgGluAsnMetIleValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleVa 367
             |||||||||||||:::|||||||||:::|||||||||||||||   ||||||   |||||
Db      1680 TCGGGAAAACATGATGGTACTAAGACAAGGATTAGAGAAAACGGGACGTTTTAAAATCGT 1739

Qy       367 lSerLysAspGluGlyValProLeuValAlaPheSerLeuLysAspSerSerCysHisTh 387
             |||||||:::   |||||||||||||||||||||||||||||||||||||||   |||
Db      1740 CTCCAAAGAAAACGGTGTTCCGTTAGTGGCGTTTTCTCTCAAAGATAGTAGCCGCCACAA 1799

Qy       387 rGluPheGluIleSerAspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMe 407
              |||||||||::::::      |||||||||:::||||||||||||||||||||||||||
Db      1800 CGAGTTCGAGGTGGCCCATACACTCCGTCGCTTCGGCTGGATCGTTCCGGCCTACACGAT 1859

Qy       407 tProProAsnAlaGlnHisIleThrValLeuArgValValIleArgGluAspPheSerAr 427
             ||||   :::|||||||||:::||||||||||||||||||||||||||||||||||||||
Db      1860 GCCTGCGGATGCGCAGCATGTCACTGTCCTTCGAGTTGTTATCCGAGAAGATTTCTCTCG 1919

Qy       427 gThrLeuAlaGluArgLeuValIleAspIleGluLysValMetArgGluLeuAspGluLe 447
             ||||||||||||||||||||||   |||   |||||||||:::   |||||||||   ||
Db      1920 AACCTTAGCCGAGAGATTGGTAGCTGATTTCGAGAAGGTTCTACACGAGCTCGATACGCT 1979

Qy       447 uProSerArgValIleHisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSerAs 467
             ||||:::||||||      |||::::::   |||:::
Db      1980 TCCGGCGAGGGTTCACGCCAAGATGGCTAATGGAAAAGTTAACGGT-------------- 2025

Qy       467 pAsnLeuMetValThrValLysLysSerAspIleAspLysGlnArgAspIleIleThrGl 487
                             |||||||||:::      :::   ||||||::::::
Db      2026 ----------------GTTAAGAAGACGCCAGAGGAGACGCAGAGAGAAGTCACGGCCTA 2069

Qy       487 yTrpLysLysPheValAlaAspArgLysLysThrSer 499
              |||||||||   :::   :::   |||||||||:::
Db      2070 CTGGAAGAAGTTGTTG---GAGACTAAGAAGACCAAC 2103

RESULT 12
AAI70004
ID   AAI70004 standard; DNA; 1180 BP.
XX
AC   AAI70004;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Tomato glutamate decarboxylase clone T-gad-19 DNA sequence.
XX
KW   Tomato; glutamate decarboxylase; GAD; transgenic plant; T-gad-19; ds.
XX
OS   Lycopersicon esculentum.
XX
PN   JP2001238555-A.
XX
PD   04-SEP-2001.
XX
PF   13-DEC-2000; 2000JP-0404172.
XX
PR   13-DEC-1999;   99JP-0352552.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 2001-610050/70.
XX
PT   Creation of a transgenic plant of improved amino acid composition,
PT   involves transformation with antisense sequences -
XX
PS   Example 1; Page 13-14; 18pp; Japanese.
XX
CC   The present invention relates to a method for creating a transgenic plant
CC   which accumulates a larger amount of at least one amino acid of edible
CC   part than a natural plant of the same species cultured under a same
CC   condition. The method involves transforming a plant with a genetic
CC   structure containing an antisense sequence of a gene sequence encoding
CC   glutamate decarboxylase (GAD) and a controlling sequence which can
```

```
C   express the antisense sequence. The resultant transgenic plant has high
C   added value. The present sequence is the DNA sequence for clone T-gad-19
C   of the tomato GAD gene, which was used to illustrate the present
C   invention.
X
Q   Sequence 1180 BP; 372 A; 188 C; 276 G; 344 T; 0 other;

lignment Scores:
red. No.:              2.32e-181      Length:          1180
core:                  1836.00        Matches:         347
ercent Similarity:     94.40%         Conservative:    24
est Local Similarity:  88.30%         Mismatches:      22
uery Match:            70.21%         Indels:          1
B:                     22             Gaps:            0

S-10-006-852-2 (1-502) x AAI70004 (1-1180)

y       38 ProLysGluAlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArg 57
           ||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
b        1 CCAAAGGAAGCAGCATATCAGATTCTAAATGATGAACTTATGTTAGATGGAAATCCAAGG 60

y       58 LeuAsnLeuAlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMet 77
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
b       61 TTGAATTTGGCATCTTTTGTGACAACATGGATGGAACCAGAATGTGACAAATTGATGATG 120

y       78 SerSerIleAsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGln 97
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      121 GATTCCATTAACAAAAATTATGTTGACATGGATGAATATCCTGTCACCACTGAGCTTCAG 180

y       98 AsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThr 117
           ||||||||||||||||||||||||||||||||||||||||||||||||:::   ||||||
b      181 AATCGGTGTGTAAACATGATAGCGCATTTATTTAATGCACCACTTGAAGATGGAGAAACT 240

y      118 AlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPhe 137
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      241 GCAGTTGGAGTTGGAACAGTTGGTTCTTCAGAAGCCATTATGCTTGCTGGATTGGCCTTT 300

y      138 LysArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIle 157
           |||||||||||||||||||||   ||||||:::|||||||||   |||||||||||||||
b      301 AAGAGAAAATGGCAAAACAAAATGAAAGCCCAAGGAAAGCCCTATGATAAGCCCAACATT 360

y      158 ValThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGlu 177
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      361 GTTACTGGTGCTAATGTCCAGGTGTGTTGGGAAAAATTTGCAAGGTATTTTGAAGTTGAG 420

y      178 LeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAsp 197
           ||||||||||||||||||::::::|||||||||||||||   |||::::::||||||:::
b      421 CTAAAAGAAGTGAAGTTGACTGATGGATACTATGTGATGAGCCCTGAGAAAGCTGTGGAA 480

y      198 MetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGlu 217
           ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
b      481 ATGGTTGATGAGAACACAATTTGTGTAGCTGCTATTTTGGGTTCAACTCTGAATGGGGAG 540

y      218 PheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrp 237
           ||||||:::||||||   |||||||||||||||   ||||||||||||||||||||||||
b      541 TTTGAGGAAGTGAAGAAATTGAATGACCTCCTATT-GAAAAGAACAAGGAAACAGGGTGG 599

y      238 AspThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrPro 257
           |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
b      600 GACACTCCAATTCATGTGGATGCAGCTAGTGGTGGATTTATTGCACCATTTATATATCCA 659

y      258 GluLeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLys 277
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      660 GAACTTGAATGGACTTTAGATTGCCATTAGTTAAAAGTATAAATGTCAGCGGTCACAAG 719

'y     278 TyrGlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuPro 297
           |||||||||||||||||||||:::||||||||||||||||||:::|||||||||||||||
b      720 TATGGCCTTGTATATGCTGGTGTCGGTTGGGTCATATGGCGGAGCAAGGAAGACTTGCCC 779

y      298 GluGluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsn 317
           :::||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
b      780 GATGAACTTATTTTTCATATTAATTACCTTGGGTCTGATCAACCTACTTTCACCCTCAAT 839

Qy     318 PheSerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHis 337
           |||||||||||||||||||||||||||||||||||||||||||||||||||   |||:::
Db     840 TTCTCTAAAGGATCAAGTCAAGTAATTGCTCAATATTATCAACTAATTCGTTTTGGTTAT 899

Qy     338 GluGlyTyrArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGly 357
           |||||||||:::|||||||||:::||||||      |||      ||||||   ||||||
Db     900 GAGGGTTATAAGAACGTCATGAAGAATTGCTTATCAAACGCAAGGGTACTAACAGAGGGA 959

Qy     358 LeuGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAla 377
           :::   |||      ||||||   |||||||||||||||   ||||||||||||||||||
Db     960 ATCACAAAAATGGGAAGATTCGAGATTGTCTCCAAGGATGTTGGTGTTCCCTTGGTTGCA 1019

Qy     378 PheSerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArg 397
           |||||||||||||||:::|||   |||   |||||||||:::|||:::   |||||||||
Db    1020 TTTTCTCTTAAAGACAATAGCAAACATGATGAGTTTGAAGTATCTGAAACTTTAAGGAGA 1079

Qy     398 TyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeu 417
           :::|||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db    1080 TTTGGATGGATTGTTCCAGCATACACTATGCCACCAGATGCTCAACACATCACAGTTCTT 1139

Qy     418 ArgValValIleArgGluAspPheSerArgThrLeuAla 430
           |||||||||||||||||||||||||||||||||||||||
Db    1140 AGAGTTGTTATTAGAGAAGATTTCTCCCGTACACTCGCC 1178

RESULT 13
AAL43412
ID   AAL43412 standard; DNA; 2493 BP.
XX
AC   AAL43412;
XX
DT   25-SEP-2002  (first entry)
XX
DE   A thaliana GAD3 coding sequence.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   P-PSDB; AAO15134.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
PT   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT   -
XX
PS   Claim 18; Page 54-55; 63pp; English.
XX
CC   The present invention relates to a method of producing a transformed
CC   plant that selectively increases production of gamma-aminobutyric acid
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the A. thaliana GAD3 coding sequence.
XX
SQ   Sequence 2493 BP; 783 A; 458 C; 507 G; 745 T; 0 other;
```

```
lignment Scores:
red. No.:               9.54e-181       Length:         2493
core:                   1834.50         Matches:        395
ercent Similarity:      52.76%          Conservative:   44
est Local Similarity:   47.48%          Mismatches:     52
uery Match:             70.15%          Indels:         342
B:                      24              Gaps:           6

S-10-006-852-2 (1-502) x AAL43412 (1-2493)

y             1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
                ||||||||||||         |||:::||||||   |||:::||||||||||||||||||
b             1 ATGGTTTTATCTAAGACAGCTTCCAAATCCGATGATTCAATCCATTCAACTTTTGCTTCC 60

y            21 ArgTyrValArgThrSerLeuPro------------------------------------ 28
                ||||||||||||   |||:::
b            61 CGTTATGTCCGCAACTCTATCTC-ACGGTAAGAAGTTGAAACACAATTTTATTTTGTTTA 119

y            28 ------------------------------------------------------------ 28

b           120 ATGTTTTCATTGGTAACTAGAGTTCTAAAACTTAGCCTAGACGACGATACACAGCATCTT 179

y            28 ------------------------------------------------------------ 28

b           180 GATTCTAGATTCAATATTTATTACAGAAATATTTATTTTTAATATACGATATAGTTCCAG 239

y            28 ------------------------------------------------------------ 28

b           240 ATTTTAATTTTTGGGTACATAAGAAAGAATACTAGATTCTAACGAAATTAACCACTTGCA 299

y            28 ------------------------------------------------------------ 28

b           300 CTGAAAGATCCGAGCATAATGTGTGTTACTATATAAGAGGTATTTTCTTTTTTAATCTTA 359

y            29 ------------------------ArgPheLysMetProGluAsnSerIleProLysGlu 40
                                        ||||||::::::|||:::||||||||||||||||||
b           360 AGCTAAATATATCAATTTTTCATCAGATTCGAAATACCTAAGAACTCGATCCCTAAGGAA 419

y            41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
                ||||||||||||||||||||||||||||||      ||||||||||||||||||||||||
b           420 GCAGCATACCAAATCATCAACGACGAGCTCAAGTTTGACGGTAACCCGAGGCTAAACCTG 479

y            61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
                |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
b           480 GCCTCCTTTGTGACCACTTGGATGGAGCCAGAATGTGACAAGCTCATGATGGAATCCATC 539

y            81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
                |||||||||   |||:::||||||:::|||||||||||||||:::|||||||||||||||
b           540 AACAAGAACAACGTTGAGATGGACCAATACCCTGTTACCACCGACCTTCAGAATCGATGC 599

y           101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
                |||||||||||||||   ||||||||||||||||||   :::   |||   |||:::|||
b           600 GTTAACATGATTGCGCGTCTCTTCAACGCGCCTTTAGGTGACGGTGAAGCCGCCATTGGT 659

y           121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
                |||||||||||||||||||||||||||:::|||||||||||||||||||||||||||:::
b           660 GTTGGCACGGTGGGGTCATCGGAGGCAGTGATGTTGGCCGGACTGGCCTTTAAGAGACAG 719

y           141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
                |||||||||||||||||||||   |||   |||   |||:::||||||||||||||||||
b           720 TGGCAGAACAAGCGTAAGGCCCTAGGGCTGCCTTATGATAGACCTAATATTGTAACCGGA 779

y           161 AlaAsn------------------------------------------------------ 162
                ||||||
b           780 GCCAATATTCAGGTAAACCAAAACAAAAATTGATTAAATTTTAAACCGGTTTAGGTCTAT 839

y           163 ---------------------------ValGlnValCysTrpGluLysPheAlaArgTyr 173
                                           :::|||||||||   ||||||||||||||||||
b           840 GTTTACATTGACTCAATTTCCGGTTCAATACAGGTTTGCTTGGAGAAATTTGCAAGGTAT 899

y           174 PheGluValGluLeuLysGluValLysLeuSerGluGlyTyrTyrValMetAspProGln 193
                ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
b           900 TTTGAAGTGGAGCTTAAGGAAGTGAAGCTGAGAGAAGGATATTACGTGATGGACCCTGAC 959

Qy          194 GlnAlaValAspMetValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThr 213
                :::||||||:::|||||||||||||||||||||||||||      |||||||||||||||
Db          960 AAAGCGGTTGAAATGGTAGACGAAAACACTATATGCGTCGTGGCCATCCTCGGTTCGACA 1019

Qy          214 LeuAsnGlyGluPheGluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLys 233
                |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1020 CTAACCGGAGAATTCGAAGACGTTAAGCTCCTCAACGACCTTTTAGTCGAGAAAAACAAG 1079

Qy          234 GluThrGly--------------------------------------------------- 236
                :::||||||
Db         1080 AAAACCGGGTAATTGAATCAAAACCAACTAACAAATTAATTTTATATACTTTTGCCTAGA 1139

Qy          236 ------------------------------------------------------------ 236

Db         1140 AATATTACAATTTCTAACGTGAGATATATTTGCTTAGAAATATTTTATTTTTTGAATGAA 1199

Qy          236 ------------------------------------------------------------ 236

Db         1200 TATAAAACTTATTAACCAAAACAAAACCATATATGTTTACATTATATGCTTCCTTGTATC 1259

Qy          236 ------------------------------------------------------------ 236

Db         1260 GAATGGTGTTTTAAATACTGATTAAAAAATGTTTTGCTTAAAAATATAACAATTTATAAT 1319

Qy          236 ------------------------------------------------------------ 236

Db         1320 GTGAGATATTCAAGCATTCTAATATCAAACCGATAAACAACAACAAACTGATTATTAATT 1379

Qy          237 ---------------------------------------TrpAspThrProIleHisVal 243
                                                       |||||||||||||||||||||
Db         1380 TATTTAACCGGTTTGGTTCCGGTTTAATATATTTGTAGATGGGATACGCCGATTCACGTG 1439

Qy          244 AspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGluTrpAspPhe 263
                ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db         1440 GACGCAGCGAGTGGTGGGTTTATTGCTCCCTTCTTGTATCCGGACTTGGAGTGGGATTTC 1499

Qy          264 ArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeuValTyrAla 283
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1500 CGGTTACCGTTGGTTAAGAGCATAAATGTGAGTGGTCACAAATACGGTTTGGTTTACGCC 1559

Qy          284 GlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeuIlePheHis 303
                |||||||||||||||:::||||||   |||   |||||||||:::|||||||||||||||
Db         1560 GGTATCGGTTGGGTCGTATGGAGAACCAAAACCGATTTGCCTGATGAACTTATCTTCCAT 1619

Qy          304 IleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys--------- 320
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1620 ATCAATTATCTTGGAGCTGATCAACCCACATTTACCCTCAACTTCTCTAAAGGTACATTA 1679

Qy          320 ------------------------------------------------------------ 320

Db         1680 CCATATCTTATGTAAAGTTTAGATATATTTATAGATTAATGTTTTGTTAATTCTTGTATA 1739

Qy          321 ------GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGlu 338
                      ||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db         1740 TTACCAGGGTCAAGTCAAGTGATTGCTCAGTACTACCAGTTGATTCGTCTTGGATTCGAG 1799

Qy          338 ------------------------------------------------------------ 338

Db         1800 GTAAATAATAACTCAATAAAGAAACTAAAACGTTACTAAATCCAATCGTATACGTACTAG 1859

Qy          338 ------------------------------------------------------------ 338

Db         1860 TATAATATACAAGTTGTTACTATACTTTATGACTACAAAAGTTCAAAACCAAGAATGTAC 1919

Qy          338 ------------------------------------------------------------ 338

Db         1920 TAAATACATTCCATAAGATTAAACGTTCCTAAATTGACAAGTTTTGGTTTTGTAGAATAG 1979

Qy          339 -------------------------GlyTyrArgAsnValMetGluAsnCysArgGluAs 350
                                         ||||||||||||||||||:::||||||||||||||
Db         1980 CTAATAATCTTTTTGTTTGGTTTAGGGATATCGCAACGTGATGGATAATTGCCGCGAGAA 2039
```

```
y         350 nMetIleValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIleValSerLysAs 370
              ||||:::|||||||||:::|||||||||||||||   |||||||||||||||||||||::
o        2040 CATGATGGTACTAAGACAAGGATTAGAGAAAACGGACGTTTTAACATCGTCTCCAAAGA 2099

y         370 pGluGlyValProLeuValAlaPheSerLeuLysAspSerSerCysHisThrGluPheGl 390
              :   |||||||||||||||||||||||||||||||||||||||   |||   ||||||||
o        2100 AAACGGTGTTCCGTTAGTGGCGTTTTCTCTCAAAGATAGTAGCCGCCACAACGAGTTCGA 2159

y         390 uIleSerAspMetLeuArgArgTyrGlyTrpIleValProAlaTyrThrMetProProAs 410
              |:::::::::||||||||||||:::||||||||||||||||||||||||||||||   ::
o        2160 GGTGGCCGAAATGCTTCGTCGCTTCGGCTGGATCGTTCCGGCCTACACGATGCCTGCGGA 2219

y         410 nAlaGlnHisIleThrValLeuArgValValIleArgGluAspPheSerArgThrLeuAl 430
              :|||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||
o        2220 TGCGCAACATGTCACGGTCCTTCGAGTTGTTATCCGAGAAGATTTCTCTCGAACCTTAGC 2279

y         430 aGluArgLeuValIleAspIleGluLysValMetArgGluLeuAspGluLeuProSerAr 450
              |||||||||||||   |||   |||||||||:::   |||||||||   ||||||:::||
o        2280 TGAGAGATTGGTAGCCGATTTCGAGAAGGTTCTACACGAGCTCGATACGCTTCCCGCGAG 2339

y         450 gValIleHisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMe 470
              ||||      |||::::::   |||:::
o        2340 GGTTCACGCCAAGATGGCTAGTGGAAAAGTTAACGGT----------------------- 2376

y         470 tValThrValLysLysSerAspIleAspLysGlnArgAspIleIleThrGlyTrpLysLy 490
                     |||||||||:::      :::   ||||||::::::         ||||||||
o        2377 -------GTTAAGAAGACGCCAGAGGAGACGCAAAGAGAAGTCACGGCCTACTGGAAGAA 2429

y         490 sPheValAlaAspArgLysLysThrSerGlyIle 501
              |||||||      :::         :::|||:::
o        2430 GTTTGTGGACACTAAGACTGACAAGAACGGCGTT 2463

ESULT 14
AL43414
D   AAL43414 standard; DNA; 1946 BP.
X
C   AAL43414;
X
T   25-SEP-2002  (first entry)
X
E   A thaliana GAD5 coding sequence.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance; gene; ds.
X
S   Arabidopsis thaliana.
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I   Kinnersley AM,  Turano FJ;
X
R   WPI; 2002-490073/52.
R   P-PSDB; AAO15136.
X
T   Making transformed plants that selectively increase gamma-aminobutyric
T   acid production, by incorporating a DNA construct with a polynucleotide
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T    -
X
S   Claim 18; Page 56; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
C   (GABA) in response to a signal, by incorporating into the plant's genome
C   a DNA construct with a non-constitutive promoter operably linked to a
```

```
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the A. thaliana GAD5 coding sequence.
XX
SQ   Sequence 1946 BP; 534 A; 382 C; 427 G; 603 T; 0 other;

Alignment Scores:
Pred. No.:              1.98e-175       Length:         1946
Score:                  1782.00         Matches:        371
Percent Similarity:     65.75%          Conservative:   57
Best Local Similarity:  56.99%          Mismatches:     62
Query Match:            68.15%          Indels:         163
DB:                     24              Gaps:           8

US-10-006-852-2 (1-502) x AAL43414 (1-1946)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              |||||||||:::         |||:::||||||      :::||||||||||||||||||
Db          1 ATGGTACTCGCA---ACCAACTCTGACTCCGACGAGCATTTGCATTCCACTTTTGCTTCT 57

Qy         21 ArgTyrValArgThrSerLeuPro------------------------------------ 28
              ||||||||||||      :::|||
Db         58 AGATATGTCCGTGCTGTTGTTCCCAGGTTCCAGAGAGTTTTGCCTCATTTTAGTTTTTTT 117

Qy         28 ------------------------------------------------------------ 28

Db        118 AATCTTGTATGCTACATTGTTATATATTTAATTATTTATGTATCTGTTTGCATATATTGA 177

Qy         29 ---ArgPheLysMetProGluAsnSerIleProLysGluAlaAlaTyrGlnIleIleAsn 47
                 |||||||||||||||::::::   :::||||||:::||||||||||||:::||||||
Db        178 AACAGGTTCAAGATGCCTGACCATTGCATGCCCAAAGATGCTGCTTATCAAGTGATCAAT 237

Qy         48 AspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeuAlaSerPheValThrThrTrp 67
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 GATGAGTTGATGCTTGATGGTAATCCCAGGCTTAACCTAGCCTCCTTTGTCACCACTTGG 297

Qy         68 MetGluProGluCysAspLysLeuIleMetSerSerIleAsnLysAsnTyrValAspMet 87
              ||||||||||||||||||||||||||||||   |||:::|||||||||||||||||||||
Db        298 ATGGAACCTGAGTGTGACAAACTCATCATGGATTCTGTCAATAAGAACTATGTTGATATG 357

Qy         88 AspGluTyrProValThrThrGluLeu--------------------------------- 96
              |||||||||||||||||||||||||||
Db        358 GATGAATATCCTGTCACCACTGAGCTCCAGGTTCCTCCTTCTTTCCTCTCATTCTCTCTC 417

Qy         97 ----------------------------------------------------------Gl 97
                                                                        ||
Db        418 TCATCTACTTTCCACTGTTTTGTCATAGACTCATACATCTTTTATCTGGCTTATTTTTCA 477

Qy         97 nAsnArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluTh 117
              |||||||||||||||||||||||||:::||||||:::||||||:::   |||   |||
Db        478 GAACCGGTGTGTAAATATGATAGCAAACTTGTTCCATGCTCCCGTTGGAGAAGACGAGGC 537

Qy        117 rAlaValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPh 137
               |||:::|||   |||||||||||||||||||||||||||||||||||||||||||||||
Db        538 TGCTATTGGGTGTGGAACTGTTGGTTCATCTGAGGCTATAATGCTTGCTGGTTTGGCTTT 597

Qy        137 eLysArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIl 157
              ||||||||||||||||::::::|||||||||:::|||   |||:::||||||||||||||
Db        598 CAAAAGGAAATGGCAACATAGGAGAAAAGCTCAGGGTCTACCTATTGATAAGCCTAACAT 657

Qy        157 eValThrGlyAlaAsnVal----------------------------------------- 163
              ||||||||||||||||
Db        658 TGTCACTGGAGCCAATGT-TCAGGTCTAAAATATTTACTTATTCTTATCCTCCAAACCAT 716

Qy        163 ------------------------------------------------------------ 163

Db        717 CACATTTGCTTTGGATAGTGATCTGTTTCTTTCCAATATCAATACATTTTCAAACTTTGT 776

Qy        164 ----------GlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGl 180
                        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        777 TTCATCCGCTCAGGTGTGCTGGGAGAAGTTTGCAAGGTACTTTGAGGTAGAGCTCAAAGA 836
```

```
y        180 uValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAs 200
             ||||||||||||||||   ||||||||||||||||||   :::||||||:::||||||||
b        837 GGTGAAACTAAGTGAAGACTACTATGTTATGGATCCAGCTAAAGCTGTAGAGATGGTGGA 896

y        200 pGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAs 220
             ||||||||||||||||||||||   ||||||||||||||||||   ||||||||||||||
b        897 TGAGAATACCATCTGTGTTGCAGCAATTCTAGGATCCACACTTACTGGAGAGTTTGAGGA 956

y        220 pValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPr 240
             |||||||   |||||||||||||||   |||||||||   ||||||||||||:::|||||
b        957 CGTTAAGCAATTGAACGATCTCTTAGCTGAGAAAAACGCAGAGACAGGATGGGAAACTCC 1016

y        240 oIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGl 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
b       1017 TATTCATGTTGATGCAGCCAGTGGAGGATTCATTGCTCCTTTCCTCTACCCTGATCTTGA 1076

y        260 uTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLe 280
             |||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
b       1077 ATGGGACTTTAGGCTTCCATGGGTGAAGAGTATTAACGTCAGTGGTCACAAGTATGGACT 1136

y        280 uValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLe 300
             |||||||||||||:::|||||||||:::||||||   |||:::|||||||||||||||||
b       1137 TGTGTATGCAGGAGTTGGTTGGGTTGTCTGGAGAACAAAAGATGATTTGCCAGAGGAACT 1196

y        300 uIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLy 320
             |:::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       1197 TGTCTTCCACATCAACTACTTGGGAGCTGATCAACCCACTTTCACTCTCAACTTCTCAAA 1256

y        320 s----------------------------------------------------------- 320
             |
b       1257 AGGTTTGTAAAATAAAAACTGGCTTTATCCAATCAAATCCATCATCACATTTCCTTTAAG 1316

y        321 ----------------------GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIl 333
                                   ||||||||||||:::||||||||||||||||||   ||
b       1317 AAACTCAATGTTTTCTTTTGCAGGGTCGAGCCAAATCATTGCTCAGTACTATCAGTTTAT 1376

y        333 eArgLeuGlyHis----------------------------------------------- 337
             ||||||||||
b       1377 CCGACTAGGCTT-TGAGGTACTTGTTCCCTTATCTGCATTACAGTTTCATTTTTTCATCT 1435

y        338 ----------------------------------GluGlyTyrArgAsnValMetGluAs 346
                                               :::||||||:::|||:::||||||||
b       1436 TGCTTAATCTAATGATTCTTTTTGGAAACTGGAAAAGGGATACAAGAACATAATGGAAAA 1495

y        346 nCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLysThrGluArgPheAsnIl 366
             ||||   :::|||         ||||||||||||:::|||   |||   :::||||||||
b       1496 CTGCATGGATAACGCAAGGAGGCTAAGAGAAGGAATAGAGATGACAGGGAAGTTCAACAT 1555

y        366 eValSerLysAspGluGlyValProLeuValAlaPheSerLeuLysAspSerSerCysHi 386
             |||||||||||||   |||||||||||||||||||||||||||||||||||||||   ||
b       1556 TGTGTCCAAAGATATTGGCGTGCCACTAGTGGCATTCTCTCTCAAAGACAGTAGCAAGCA 1615

y        386 sThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrpIleValProAlaTyrTh 406
             ||||   |||||||||::::::   ||||||::::::|||||||||:::|||||||||||
b       1616 CACGGTGTTTGAGATCGCAGAGTCTTTGAGAAAATTCGGGTGGATCATACCGGCTTACAC 1675

y        406 rMetProProAsnAlaGlnHisIleThrValLeuArgValValIleArgGluAspPheSe 426
             |||||||   :::||||||||||||   ||||||||||||||||||||||||||||||||
b       1676 TATGCCTGCAGATGCACAGCACATTGCTGTGCTCAGAGTTGTGATAAGAGAAGACTTTAG 1735

y        426 rArgThrLeuAlaGluArgLeuValIleAspIleGluLysValMetArgGluLeuAspGl 446
             ||||   ||||||:::||||||:::      |||   :::|||::::::|||::::::
b       1736 CCGAGGCCTTGCAGATAGACTCATCACACATATCATTCAGGTGCTGAAAGAGATTGAAGG 1795

y        446 uLeuProSerArgValIleHisLysIleSerLeuGlyGlnGluLysSerGluSerAsnSe 466
              ||||||||||||:::   |||      :::               :::      :::
b       1796 GCTTCCTAGCAGGATTGCACATCTTGCTGCG------------GCTGCAGCGGTTAGTGG 1843

y        466 rAspAsnLeuMetValThrValLysLysSerAspIleAspLysGlnArgAspIleIleTh 486
              |||:::      |||   ||||||   :::   :::      :::   ||||||
b       1844 TGATGATGAAGAAGTTAAAGTGAAGACTGCCAAGATGTCCTTGGAG---GATATCACTAA 1900
```

```
Qy         486 rGlyTrpLysLysPheValAlaAspArgLys 496
                   ||||||:::   |||      ::::::
Db        1901 GTATTGGAAACGCCTTGTGGAACACAAGAGA 1931

RESULT 15
AAC44889
ID   AAC44889 standard; DNA; 1607 BP.
XX
AC   AAC44889;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 44511.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
```

```
  18-JUN-1999;   99US-0139455.
  18-JUN-1999;   99US-0139456.
  18-JUN-1999;   99US-0139457.
R 18-JUN-1999;   99US-0139458.
R 18-JUN-1999;   99US-0139459.
R 18-JUN-1999;   99US-0139460.
R 18-JUN-1999;   99US-0139461.
R 18-JUN-1999;   99US-0139462.
R 18-JUN-1999;   99US-0139463.
R 18-JUN-1999;   99US-0139750.
R 18-JUN-1999;   99US-0139763.
R 21-JUN-1999;   99US-0139817.
R 22-JUN-1999;   99US-0139899.
R 23-JUN-1999;   99US-0140353.
R 23-JUN-1999;   99US-0140354.
R 24-JUN-1999;   99US-0140695.
R 28-JUN-1999;   99US-0140823.
R 29-JUN-1999;   99US-0140991.
R 30-JUN-1999;   99US-0141287.
R 01-JUL-1999;   99US-0141842.
R 01-JUL-1999;   99US-0142154.
R 02-JUL-1999;   99US-0142055.
R 06-JUL-1999;   99US-0142390.
R 08-JUL-1999;   99US-0142803.
R 09-JUL-1999;   99US-0142920.
R 12-JUL-1999;   99US-0142977.
R 13-JUL-1999;   99US-0143542.
R 14-JUL-1999;   99US-0143624.
R 15-JUL-1999;   99US-0144005.
R 16-JUL-1999;   99US-0144085.
R 16-JUL-1999;   99US-0144086.
R 19-JUL-1999;   99US-0144325.
R 19-JUL-1999;   99US-0144331.
R 19-JUL-1999;   99US-0144332.
R 19-JUL-1999;   99US-0144333.
R 19-JUL-1999;   99US-0144334.
R 19-JUL-1999;   99US-0144335.
R 20-JUL-1999;   99US-0144352.
R 20-JUL-1999;   99US-0144632.
R 20-JUL-1999;   99US-0144884.
R 21-JUL-1999;   99US-0144814.
R 21-JUL-1999;   99US-0145086.
R 21-JUL-1999;   99US-0145088.
R 22-JUL-1999;   99US-0145085.
R 22-JUL-1999;   99US-0145087.
R 22-JUL-1999;   99US-0145089.
R 22-JUL-1999;   99US-0145192.
R 23-JUL-1999;   99US-0145145.
R 23-JUL-1999;   99US-0145218.
R 23-JUL-1999;   99US-0145224.
R 26-JUL-1999;   99US-0145276.
R 27-JUL-1999;   99US-0145913.
R 27-JUL-1999;   99US-0145918.
R 27-JUL-1999;   99US-0145919.
R 28-JUL-1999;   99US-0145951.
R 02-AUG-1999;   99US-0146386.
R 02-AUG-1999;   99US-0146388.
R 02-AUG-1999;   99US-0146389.
R 03-AUG-1999;   99US-0147038.
R 04-AUG-1999;   99US-0147204.
R 04-AUG-1999;   99US-0147302.
R 05-AUG-1999;   99US-0147192.
R 05-AUG-1999;   99US-0147260.
R 06-AUG-1999;   99US-0147303.
R 06-AUG-1999;   99US-0147416.
R 09-AUG-1999;   99US-0147493.
R 09-AUG-1999;   99US-0147935.
R 10-AUG-1999;   99US-0148171.
R 11-AUG-1999;   99US-0148319.
R 12-AUG-1999;   99US-0148341.
R 13-AUG-1999;   99US-0148565.
R 13-AUG-1999;   99US-0148684.
R 16-AUG-1999;   99US-0149368.
```

```
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

Alignment Scores:
Pred. No.:              1.85e-153      Length:         1607
Score:                  1570.00        Matches:        337
Percent Similarity:     69.64%         Conservative:   46
Best Local Similarity:  61.27%         Mismatches:     51
Query Match:            60.04%         Indels:         117
DB:                     21             Gaps:           10

US-10-006-852-2 (1-502) x AAC44889 (1-1607)

Qy         1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
             |||||||||:::         :::   :::|||   ||||||   :::   |||   |||
```

```
  1 ATGGTTTTGACAAAAACCGCAACG---AATGATGAATCTGTCTGCACCATGTTCGGATCT 57

 21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
    |||||||||||||||:::||||||::::::::::::   ||||||||||||||||||:::
 58 CGCTATGTTCGCACTACACTTCCCAAGTATGAGATTGGTGAGAATTCGATACCGAAAGAC 117

 41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
    ||||||||||||||||||   ||||||||||||||||||
118 GCTGCGTATCAGATCATAAAAGATGAGCTGATGCTTGAT--------------------- 156

 61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
                                  |||||||||||||||||||||   ||||||
157 ------------------------------GAGTGTGACAAACTCATCATGGACTCTATC 186

 81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGln--------- 97
    |||   |||||||||||||||||||||||||||||||||||||||||||||
187 AAT---AACTACGTTGATATGGATGAGTACCCTGTCACAACTGAGCTCCAGTTGTATGAA 243

 97 ------------------------------------------------------------ 97

244 AGTTCATGTGGCGTAGTGGTAAATACTCGCAAATGCATTCTCTCCACCCAGGGGTATAAT 303

 98 ---------------------------------------------------------Asn 98
                                                             |||
304 CGTAAATTAAACACAACTTCTTCTCCCAAACAAAACCCTAGTAGTCGCCGTTCCTCTAAC 363

 99 ArgCysValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAla 118
    ||||||||||||:::||||||   ||||||||||||||||||||||||:::|||||||||
364 CGATGTGTAAACATTATAGCTCGACTGTTCAATGCGCCACTCGAGGAATCTGAGACGGCG 423

119 ValGlyValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLys 138
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
424 GTGGGAGTAGGGACAGTTGGTTCTTCAGAAGCCATCATGTTAGCCGGATTGGCCTTCAAA 483

139 ArgLysTrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleVal 158
    |||||||||||||||||||||||||||||||||||||||   |||||||||||||||
484 AGAAAATGGCAGAACAAACGCAAGGCTGAGGGTAAACCCTATGACAAACCCAACATTTGT 543

159 ThrGlyAlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeu 178
                ::::::      |||||||||||||||||||||||||||||||||
544 CACCGGATGCCAATCAAGTTTGCTTGGGAGAAATTCGCTCGGTACTTCGAGGTGGAG--- 600

179 LysGluValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMet 198
    :::   |||   ||||||||||||||||||||||||||||||   :::|||   :::|||
601 CAAGGAGTAAACCTAAGTGAAGGTTACTACGTGATGGATCCAGACAAAGCAGCAGAAATG 660

199 ValAspGluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPhe 218
    |||||||||||||||||||||||||||   |||||||||||||||||||||
661 GTAGACGAGAACACAATCTGTGTCGCA---ATATTGGGATCCACACTCAAC--------- 708

219 GluAspValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAsp 238

708 ------------------------------------------------------------ 708

239 ThrProIleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGlu 258
                            :::   |||   ||||||||||||:::|||||||||
709 ------------------------GCAAGTGGGGTCATAGCTCCGTTTATCTATCCTGAA 744

259 LeuGluTrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyr 278
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
745 TTAGAATGGGACTTTAGGCTTCCTTTGGTTAAGAGCATCAACGTGAGTGGTCACAAGTAT 804

279 GlyLeuValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGlu 298
    ||||||||||||||||||||||||||||||:::||||||      |||||||||||||||
805 GGACTAGTCTATGCTGGTATTGGTTGGGTCGTGTGGAGGGCAGCAGAGGATTTACCTGAA 864

299 GluLeuIlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPhe 318
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
865 GAGCTTATCTTTCATATTAATTATCTTGGTGCTGATCAACCCACTTTCACTCTCAATTTC 924

319 SerLysGlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIle-ArgLeuGlyHisGl 338
    |||   ||||||||||||:::|||||||||||||||||||||||| |||||||||
925 TCCAA-GGATCGAGCCAAATTATTGCTCAATACTACCAGCTCATTTCGTCTTGGA----- 978

Qy     338 uGlyTyrArgAsnValMetGluAsnCysArgGlu-AsnMetIleValLeuArgGluGlyL 358
                              |||||||||   ||| ||||||:::||||||:::||||||:
Db     979 ----------------TTCGAGAATTGCATAGAAGAACATGGTGGTTCTCAAAGAAGGTA 1022

Qy     358 euGluLysThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaP 378
           ::||||||||||||||||||||||||||||||||||||:::|||||||||:::|||||||
Db    1023 TAGAGAAAACAGAGCGTTTCAACATAGTCTCAAAGGACCAAGGAGTGCCAGTCGTCGCCT 1082

Qy     378 he-SerLeuLysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArg 397
           || ||||||||||||   |||   |||   |||||||||||||||:::||||||||||||
Db    1083 TCTTCTCTCAAGGACCATAGTTTCCACAACGAGTTCGAGATCTCTGAGATGCTACGTCGT 1142

Qy     398 TyrGlyTrpIleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeu 417
           :::||||||||||||||||||||||||||||||   :::|||||||||||||||||||||
Db    1143 TTTGGCTGGATCGTCCCAGCTTACACTATGCCTGCCGATGCACAGCACATCACGGTTCTG 1202

Qy     418 ArgValValIleArgGluAspPheSerArgThrLeuAlaGlu-ArgLeuValIleAspIl 437
           |||||||||||||||||||||||||||||||||||||||||| |||||||||   |||||
Db    1203 CGTGTTGTCATCAGGGAAGATTTCTCAAGAACACTCGCGGAAGAGACTTGTTGCTGATAT 1262

Qy     437 eGluLysValMetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLe 457
           |   ||||||:::   |||||||||   |||||||||::::::      |||:::   ::
Db    1263 TTCGAAGGTGCTTCATGAGCTAGATACCTTGCCTTCCAAGATATCTAAGAAGATGGGAAT 1322

Qy     457 uGlyGlnGluLysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAs 477
           :                           :::   :::         ||||||:::
Db    1323 A---------------------------GAAGGGATCGCGGAAAATGTAAAGGAGAAGAA 1355

Qy     477 pIleAspLysGln------ArgAspIleIleThrGlyTrpLysLysPheValAlaAspAr 495
            ::::::|||:::         ::::::|||   ||||||:::|||||||||   :::||
Db    1356 GATGGAGAAGGAGATTCTGATGGAAGTTATTGTTGGATGGAGGAAGTTTGTGAAGGAGAG 1415

Qy     495 gLysLysThrSerGlyIleCys 502
           |||||||   :::|||:::|||
Db    1416 GAAGAAGATGAATGGTGTGTGC 1437

Search completed: October 22, 2003, 15:31:10
Job time : 406 secs
```

```
                    GenCore version 5.1.6
               Copyright (c) 1993 - 2003  Compugen Ltd.

M protein - nucleic search, using frame_plus_p2n model

un on:         October 22, 2003, 15:09:39 ; Search time 4557 Seconds
                                            (without alignments)
                                            4506.612 Million cell updates/sec

itle:          US-10-006-852-2
erfect score:  2615
equence:       1 MVLSHAVSESDVSVHSTFAS..........DIITGWKKFVADRKKTSGIC 502

coring table:  BLOSUM62
               Xgapop 10.0 , Xgapext  0.5
               Ygapop 10.0 , Ygapext  0.5
               Fgapop  6.0 , Fgapext  7.0
               Delop   6.0 , Delext   7.0

earched:       2888711 seqs, 20454813386 residues

otal number of hits satisfying chosen parameters:      5777422

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

ommand line parameters:
MODEL=frame+_p2n.model -DEV=xlp
Q=/cgn2_1/USPTO_spool_p/US10006852/runat_21102003_164103_2453/app_query.fasta_1.647
DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US10006852_@CGN_1_1_4958_@runat_21102003_164103_2453 -NCPU=6 -ICPU=3
NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase :      GenEmbl:*
               1:  gb_ba:*
               2:  gb_htg:*
               3:  gb_in:*
               4:  gb_om:*
               5:  gb_ov:*
               6:  gb_pat:*
               7:  gb_ph:*
               8:  gb_pl:*
               9:  gb_pr:*
               10:  gb_ro:*
               11:  gb_sts:*
               12:  gb_sy:*
               13:  gb_un:*
               14:  gb_vi:*
               15:  em_ba:*
               16:  em_fun:*
               17:  em_hum:*
               18:  em_in:*
               19:  em_mu:*
               20:  em_om:*
               21:  em_or:*
               22:  em_ov:*
               23:  em_pat:*
               24:  em_ph:*
               25:  em_pl:*
               26:  em_ro:*
               27:  em_sts:*
               28:  em_un:*
               29:  em_vi:*
               30:  em_htg_hum:*
               31:  em_htg_inv:*
               32:  em_htg_other:*
               33:  em_htg_mus:*
               34:  em_htg_pln:*
               35:  em_htg_rod:*
               36:  em_htg_mam:*
               37:  em_htg_vrt:*
               38:  em_sy:*
               39:  em_htgo_hum:*
               40:  em_htgo_mus:*
               41:  em_htgo_other:*

  Pred. No. is the number of results predicted by chance to have a
  score greater than or equal to the score of the result being printed,
  and is derived by analysis of the total score distribution.

                              SUMMARIES
                       %
Result              Query
  No.     Score     Match Length DB  ID                Description
---------------------------------------------------------------------------
    1      2615     100.0   1509  8  ATU10034          U10034 Arabidopsis
    2      2607      99.7   1509  6  AX507115          AX507115 Sequence
    3      2607      99.7   1509  6  AX651552          AX651552 Sequence
    4      2607      99.7   1509  8  BT001047          BT001047 Arabidops
    5      2607      99.7   1813  8  AY094464          AY094464 Arabidops
    6      2250      86.0   1785  8  PETGADX           L16797 Petunia hyb
    7      2250      86.0   1785  8  PETGLUDECA        L16977 Petunia hyb
    8      2240      85.7   1705  8  AF020425          AF020425 Nicotiana
    9      2233      85.4   1745  8  AF506366          AF506366 Nicotiana
   10      2232      85.4   1929  8  NTU54774          U54774 Nicotiana t
   11      2228      85.2   1672  8  AF352732          AF352732 Nicotiana
   12      2200      84.1   1479  6  AX654603          AX654603 Sequence
   13      2200      84.1   1502  8  AY187941          AY187941 Oryza sat
   14      2199      84.1   1771  8  AF020424          AF020424 Nicotiana
   15    2186.5      83.6   1776  8  AF353615          AF353615 Nicotiana
   16    2135.5      81.7   1482  6  AX412801          AX412801 Sequence
   17    2135.5      81.7   1482  6  AX506242          AX506242 Sequence
   18    2135.5      81.7   1482  6  AX651853          AX651853 Sequence
   19    2135.5      81.7   1482  8  AY124860          AY124860 Arabidops
   20    2135.5      81.7   1792  8  AF361836          AF361836 Arabidops
   21      2130      81.5   1485  8  AY124873          AY124873 Arabidops
   22      2130      81.5   1676  8  ATU46665          U46665 Arabidopsis
   23      2130      81.5   1681  8  ATU49937          U49937 Arabidopsis
   24      2130      81.5   1695  8  AF428372          AF428372 Arabidops
   25      2126      81.3   1685  8  AF428294          AF428294 Arabidops
   26      2102      80.4   1650  8  AY081259          AY081259 Arabidops
   27      2080      79.5   1712  8  AK118125          AK118125 Arabidops
   28    2065.5      79.0  75188  8  AB005238          AB005238 Arabidops
   29    2056.5      78.6   1779  8  AB056060          AB056060 Oryza sat
   30      2014      77.0   1783  6  A39964            A39964 Sequence 8
   31      2014      77.0   1783  6  AR070945          AR070945 Sequence
   32      2014      77.0   1783  8  LEGDL             X80840 L.esculentu
   33      2007      76.7   1455  6  AX653206          AX653206 Sequence
   34      1965      75.1   1605  8  AY084689          AY084689 Arabidops
c  35      1945      74.4 136098  8  AC137696          AC137696 Genomic s
c  36    1936.5      74.1  76976  8  AC006532          AC006532 Arabidops
   37      1894      72.4 147227  8  OSJN00013         AL606448 Oryza sat
   38      1854      70.9   1467  6  AX653024          AX653024 Sequence
   39      1844      70.5   1781  8  AB056061          AB056061 Oryza sat
   40      1836      70.2   1180  6  BD017613          BD017613 Method fo
c  41      1782      68.1  27194  8  AB026646          AB026646 Arabidops
c  42      1699      65.0 147227  8  OSJN00013         AL606448 Oryza sat
   43      1696      64.9   3983  8  AB056063          AB056063 Oryza sat
   44      1630      62.3   5160  6  AX461514          AX461514 Sequence
   45      1630      62.3  81490  8  F12P19            AC009513 Sequence

                              ALIGNMENTS

RESULT 1
```

```
TU10034
OCUS       ATU10034                1509 bp    mRNA    linear   PLN 30-MAR-1996
EFINITION  Arabidopsis thaliana Columbia glutamate decarboxylase (GAD) mRNA,
           complete cds.
CCESSION   U10034
ERSION     U10034.1  GI:497978
EYWORDS    .
OURCE      Arabidopsis thaliana (thale cress)
 ORGANISM  Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
EFERENCE   1  (bases 1 to 1509)
 AUTHORS   Arazi,T., Baum,G., Snedden,W.A., Shelp,B.J. and Fromm,H.
 TITLE     Molecular and biochemical analysis of calmodulin interactions with
           the calmodulin-binding domain of plant glutamate decarboxylase
 JOURNAL   Plant Physiol. 108 (2), 551-561 (1995)
 MEDLINE   95334488
  PUBMED   7610159
EFERENCE   2  (bases 1 to 1509)
 AUTHORS   Fromm,H.
 TITLE     Direct Submission
 JOURNAL   Submitted (25-MAY-1994) Hillel Fromm, Department of Plant Genetics,
           Weizmann Institute of Science, Rehovot 76100, Israel
EATURES             Location/Qualifiers
    source          1. .1509
                    /organism="Arabidopsis thaliana"
                    /mol_type="mRNA"
                    /strain="Columbia"
                    /db_xref="taxon:3702"
                    /clone="35B3T7"
    gene            1. .1509
                    /gene="GAD"
    CDS             1. .1509
                    /gene="GAD"
                    /EC_number="4.1.1.15"
                    /note="GAD1 isoform; similar to the GAD2 isoform of
                    Arabidopsis thaliana, GenBank Accession Number U49937"
                    /codon_start=1
                    /product="glutamate decarboxylase"
                    /protein_id="AAA93132.1"
                    /db_xref="GI:497979"
                    /translation="MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                    IINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVN
                    MIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTG
                    ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEF
                    EDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                    KYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                    LGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEIS
                    DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSR
                    VIHKISLGQEKSESNSDNLMVTVKKSDIDKQRDIITGWKKFVADRKKTSGIC"
    misc_feature    1417. .1506
                    /gene="GAD"
                    /note="calmodulin-binding domain"
ASE COUNT      433 a    321 c    385 g    370 t
RIGIN

lignment Scores:
red. No.:              3.42e-212      Length:         1509
core:                  2615.00        Matches:        502
ercent Similarity:     100.00%        Conservative:   0
est Local Similarity:  100.00%        Mismatches:     0
uery Match:            100.00%        Indels:         0
B:                     8              Gaps:           0

S-10-006-852-2 (1-502) x ATU10034 (1-1509)

y          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

y         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

Qy          41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

Qy          61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

Qy          81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

Qy         101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

Qy         121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy         141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy         161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy         181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy         201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 GAGAACACCATTTGTGTTGCGGACATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy         221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy         241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy         261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy         281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy         301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy         321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy         341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy         361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy         381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200
```

```
       401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

       421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

       441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

y      461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

y      481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

y      501 IleCys 502
           ||||||
      1501 ATCTGC 1506

SULT 2
507115
CUS         AX507115                1509 bp    DNA     linear   PAT 27-SEP-2002
CFINITION   Sequence 1810 from Patent WO0216655.
CCESSION    AX507115
ERSION      AX507115.1  GI:23388352
CYWORDS     .
OURCE       Arabidopsis thaliana (thale cress)
 ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
CFERENCE    1
 AUTHORS    Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
 TITLE      Stress-regulated genes of plants, transgenic plants containing
            same, and methods of use
 JOURNAL    Patent: WO 0216655-A 1810 28-FEB-2002;
            The Scripps Research Institute (US) ; Syngenta Participations AG
            (CH)
CATURES              Location/Qualifiers
    source           1. .1509
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3702"
ASE COUNT      432 a    322 c    385 g    370 t
RIGIN

lignment Scores:
red. No.:               1.63e-211      Length:         1509
core:                   2607.00        Matches:        501
ercent Similarity:      99.80%         Conservative:   0
est Local Similarity:   99.80%         Mismatches:     1
uery Match:             99.69%         Indels:         0
B:                      6              Gaps:           0

S-10-006-852-2 (1-502) x AX507115 (1-1509)

y        1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

y       21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

y       41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180
```

```
Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260
```

```
y        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

y        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

y        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

y        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

y        501 IleCys 502
             ||||||
o       1501 ATCTGC 1506

ESULT 3
X651552
OCUS       AX651552                 1509 bp    DNA     linear   PAT 22-MAR-2003
EFINITION  Sequence 360 from Patent WO03000898.
CCESSION   AX651552
ERSION     AX651552.1  GI:29154370
EYWORDS    .
OURCE      Arabidopsis thaliana (thale cress)
 ORGANISM  Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
EFERENCE   1
 AUTHORS   Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
           Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
 TITLE     Plant genes involved in defense against pathogens
 JOURNAL   Patent: WO 03000898-A 360 03-JAN-2003;
           Syngenta Participations AG (CH)
EATURES             Location/Qualifiers
    source          1..1509
                    /organism="Arabidopsis thaliana"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:3702"
ASE COUNT      432 a    322 c    385 g    370 t
RIGIN

lignment Scores:
red. No.:              1.63e-211      Length:        1509
core:                  2607.00        Matches:       501
ercent Similarity:     99.80%         Conservative:  0
est Local Similarity:  99.80%         Mismatches:    1
uery Match:            99.69%         Indels:        0
B:                     6              Gaps:          0

S-10-006-852-2 (1-502) x AX651552 (1-1509)

y          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

y         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

y         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

y         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

y         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

Qy       101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

Qy       121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy       141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy       161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy       181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600

Qy       201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
             |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

Qy       221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

Qy       241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

Qy       261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

Qy       281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

Qy       301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

Qy       321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

Qy       341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

Qy       361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

Qy       381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

Qy       401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

Qy       421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

Qy       441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
o         1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

y          461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o         1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

y          481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
o         1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

y          501 IleCys 502
               ||||||
o         1501 ATCTGC 1506

ESULT 4
T001047
OCUS       BT001047                1509 bp    mRNA    linear   PLN 18-OCT-2002
EFINITION  Arabidopsis thaliana At5g17330/MKP11_18 mRNA, complete cds.
CCESSION   BT001047
ERSION     BT001047.1  GI:24111354
EYWORDS    FLI_CDNA.
OURCE      Arabidopsis thaliana (thale cress)
 ORGANISM  Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
EFERENCE   1  (bases 1 to 1509)
 AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
           Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
           Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
           Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
           Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
           Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yamada,K.,
           Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
           Ecker,J.R.
 TITLE     Arabidopsis ORF clones
 JOURNAL   Unpublished
EFERENCE   2  (bases 1 to 1509)
 AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
           Chan,M.M., Chang,C.H., Dale,J.M., Hayashizaki,Y., Hsuan,V.W.,
           Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
           Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
           Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A.,
           Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yamada,K.,
           Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and
           Ecker,J.R.
 TITLE     Direct Submission
 JOURNAL   Submitted (18-OCT-2002) Salk Institute Genomic Analysis Laboratory
           (SIGnAL), Plant Biology Laboratory, The Salk Institute for
           Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
           USA
OMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
           collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
           Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
           Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
           Hayashizaki,Y. and Shinozaki,K.

           The Salk, Stanford, PGEC (SSP) Consortium members constructed and
           sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
           Chen,H., Kim,C.J., Shinn,P., Banh,J. Bowser,L., Chan,M.M.,
           Chang,E., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G.,
           Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S.,
           Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M.,
           Wu,H.C., Yamada,K., Yamamura, Y., Yu,G., Davis,R.W.,  Theologis,A.,
           and Ecker,J.R.

           Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
           this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
           contributed equally to this work as PIs.
EATURES             Location/Qualifiers
   source           1. .1509
                    /organism="Arabidopsis thaliana"
                    /mol_type="mRNA"
                    /db_xref="taxon:3702"
                    /chromosome="5"
                    /clone="U21967"
                    /note="This clone is in pUNI 51
                    ecotype: Columbia"
     CDS            1. .1509
                    /note="glutamate decarboxylase 1 (GAD 1) (sp|Q42521)"
                    /codon_start=1
                    /product="At5g17330/MKP11_18"
                    /protein_id="AAN46801.1"
                    /db_xref="GI:24111355"
                    /translation="MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                    IINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVN
                    MIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTG
                    ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVAAILGSTLNGEF
                    EDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                    KYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                    LGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEIS
                    DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSR
                    VIHKISLGQEKSESNSDNLMVTVKKSDIDKQRDIITGWKKFVADRKKTSGIC"
BASE COUNT      432 a    322 c    385 g    370 t
ORIGIN

Alignment Scores:
Pred. No.:              1.63e-211       Length:          1509
Score:                  2607.00         Matches:         501
Percent Similarity:     99.80%          Conservative:    0
Best Local Similarity:  99.80%          Mismatches:      1
Query Match:            99.69%          Indels:          0
DB:                     8               Gaps:            0

US-10-006-852-2 (1-502) x BT001047 (1-1509)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 60

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 120

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 180

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 240

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 300

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 360

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 420

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 480

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 540

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 600
```

```
y        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
             |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
ɔ        601 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 660

y        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ        661 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 720

y        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ        721 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 780

y        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        781 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 840

y        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ        841 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 900

y        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ        901 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 960

y        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ        961 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1020

y        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ       1021 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1080

y        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ       1081 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1140

y        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ       1141 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1200

y        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ       1201 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1260

y        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ       1261 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1320

y        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ       1321 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1380

y        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ       1381 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1440

y        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ɔ       1441 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1500

y        501 IleCys 502
             ||||||
ɔ       1501 ATCTGC 1506

ESULT 5
Y094464
OCUS       AY094464                 1813 bp    mRNA    linear   PLN 05-MAY-2002
EFINITION  Arabidopsis thaliana AT5g17330/MKP11_18 mRNA, complete cds.
CCESSION   AY094464
ERSION     AY094464.1  GI:20453186
EYWORDS    FLI_CDNA.
OURCE      Arabidopsis thaliana (thale cress)
 ORGANISM  Arabidopsis thaliana
```

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1813)
  AUTHORS   Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
            Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
            Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
            Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
  TITLE     Arabidopsis cDNA clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1813)
  AUTHORS   Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,
            Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
            Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
            Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
            Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
            Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H.,
            Cheuk,R., Kim,C.J.,  Meyers,M.C., Banh,J. Bowser,L., Chang,E.,
            Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B.,
            Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
            Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
            Yamada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W.,  Theologis,A.,
            and Ecker,J.R.

            Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1813
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="RAFL09-48-C10(R21967)"
                     /note="ecotype: Columbia"
     CDS             74. .1582
                     /note="glutamate decarboxylase 1 (GAD 1) (sp|Q42521)"
                     /codon_start=1
                     /product="AT5g17330/MKP11_18"
                     /protein_id="AAM19834.1"
                     /db_xref="GI:20453187"
                     /translation="MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVAAILGSTLNGEF
                     EDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEIS
                     DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSR
                     VIHKISLGQEKSESNSDNLMVTVKKSDIDKQRDIITGWKKFVADRKKTSGIC"
BASE COUNT      504 a    384 c    417 g    508 t
ORIGIN

Alignment Scores:
```

```
red. No.:               2.04e-211      Length:        1813
core:                   2607.00        Matches:       501
ercent Similarity:      99.80%         Conservative:  0
est Local Similarity:   99.80%         Mismatches:    1
uery Match:             99.69%         Indels:        0
B:                      8              Gaps:          0

S-10-006-852-2 (1-502) x AY094464 (1-1813)

y           1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b          74 ATGGTGCTCTCCCACGCCGTATCGGAGTCGGACGTCTCCGTCCACTCCACATTCGCATCA 133

y          21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         134 CGTTACGTCCGTACTTCACTTCCTAGGTTCAAGATGCCGGAAAACTCGATTCCTAAGGAA 193

y          41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         194 GCGGCGTATCAGATCATCAACGACGAGCTGATGCTTGACGGGAATCCACGGTTGAACTTA 253

y          61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         254 GCCTCCTTTGTGACGACATGGATGGAGCCTGAGTGTGATAAACTCATCATGTCCTCCATC 313

y          81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         314 AACAAGAACTATGTTGACATGGACGAGTACCCCGTCACCACCGAACTTCAGAACCGATGT 373

y         101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         374 GTGAACATGATTGCACATCTATTCAATGCACCGTTAGAAGAGGCGGAGACCGCCGTCGGA 433

y         121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         434 GTAGGAACCGTTGGATCATCGGAGGCCATAATGTTGGCCGGTTTGGCCTTCAAGCGTAAA 493

y         141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         494 TGGCAGAACAAGCGCAAAGCTGAAGGCAAACCCGTCGATAAACCCAACATTGTCACCGGA 553

y         161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         554 GCCAATGTTCAAGTGTGTTGGGAGAAATTCGCTAGGTACTTTGAGGTTGAACTTAAGGAA 613

y         181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         614 GTGAAATTGAGTGAAGGATACTATGTGATGGACCCTCAACAAGCTGTTGATATGGTTGAT 673

y         201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
b         674 GAGAACACCATTTGTGTTGCGGCCATTCTTGGTTCCACTCTTAATGGAGAATTCGAAGAT 733

y         221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         734 GTTAAACTCTTGAACGATCTCTTGGTCGAAAAGAACAAAGAAACCGGATGGGATACACCA 793

y         241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         794 ATCCACGTGGATGCGGCAAGTGGAGGATTCATTGCACCGTTTTTGTATCCGGAATTGGAA 853

y         261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         854 TGGGACTTTAGACTTCCCTTGGTGAAGAGTATCAATGTGAGTGGTCACAAGTATGGACTT 913

y         281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         914 GTGTACGCAGGGATTGGTTGGGTGATCTGGAGAAACAAAGAGGATTTGCCTGAGGAACTC 973

y         301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         974 ATCTTCCATATCAATTATCTTGGTGCTGACCAACCCACCTTTACTCTCAATTTCTCCAAA 1033
```

```
Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1034 GGTTCAAGTCAAGTCATTGCTCAATACTACCAACTTATCCGATTGGGCCACGAGGGTTAC 1093

Qy        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1094 AGAAATGTGATGGAGAATTGCAGAGAGAATATGATCGTCCTAAGGGAAGGACTTGAGAAG 1153

Qy        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1154 ACAGAAAGGTTCAACATCGTCTCAAAGGACGAGGGAGTGCCACTTGTCGCTTTCTCCTTG 1213

Qy        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1214 AAAGATAGCAGCTGTCACACTGAGTTCGAAATCTCCGACATGCTTCGCAGGTATGGATGG 1273

Qy        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 ATAGTGCCGGCCTACACAATGCCTCCAAATGCACAACACATCACTGTTCTTCGTGTGGTT 1333

Qy        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 ATCAGAGAAGATTTCTCGAGAACACTCGCTGAGAGACTTGTGATCGATATAGAGAAAGTG 1393

Qy        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1394 ATGCGTGAGCTCGATGAGCTTCCTTCGAGAGTGATTCACAAAATATCACTTGGACAAGAG 1453

Qy        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 AAGAGTGAATCTAACAGCGATAACTTGATGGTCACGGTGAAGAAGAGCGATATCGACAAG 1513

Qy        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1514 CAGAGAGATATCATCACTGGCTGGAAGAAGTTTGTCGCCGACAGGAAGAAGACGAGTGGT 1573

Qy        501 IleCys 502
              ||||||
Db       1574 ATCTGC 1579

RESULT 6
PETGADX
LOCUS       PETGADX                 1785 bp    mRNA    linear   PLN 15-OCT-1993
DEFINITION  Petunia hybrida glutamate decarboxylase (gad) mRNA, complete cds.
ACCESSION   L16797
VERSION     L16797.1  GI:294111
KEYWORDS    glutamate decarboxylase.
SOURCE      Petunia x hybrida
  ORGANISM  Petunia x hybrida
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Petunia.
REFERENCE   1  (bases 1 to 1785)
  AUTHORS   Baum,G., Chen,Y., Arazi,T., Takatsuji,H. and Fromm,H.
  TITLE     A plant glutamate decarboxylase containing a calmodulin binding
            domain. Cloning, sequence, and functional analysis
  JOURNAL   J. Biol. Chem. 268 (26), 19610-19617 (1993)
  MEDLINE   93374956
   PUBMED   8366104
COMMENT     Original source text: Petunia hybrida flower corolla cDNA to mRNA.
FEATURES             Location/Qualifiers
     source          1. .1785
                     /organism="Petunia x hybrida"
                     /mol_type="mRNA"
                     /db_xref="taxon:4102"
                     /tissue_type="corolla"
                     /dev_stage="flower"
     gene            1. .1785
                     /gene="gad"
     CDS             72. .1574
                     /gene="gad"
                     /EC_number="4.1.1.15"
                     /codon_start=1
```

```
                     /product="glutamate decarboxylase"
                     /protein_id="AAA33709.1"
                     /db_xref="GI:294112"
                     /translation="MVLSKTVSQSDVSIHSTFASRYVRTSLPRFKMPDNSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLEDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLVEKNKETGWDTPIHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWVVWRNKDDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGYEGYKNVMENCQENASVLREGLEKTGRFNIISKEIGVPLVAFSLKDNRQHNEFEIS
                     ETLRRFGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVRDIEKVLHELDTLPAR
                     VNAKLAVAEEQAAANGSEVHKKTDSEVQLEMITAWKKFVEEKKKKTNRVC"
ASE COUNT      550 a    298 c    401 g    536 t
RIGIN

lignment Scores:
red. No.:              3.5e-181       Length:          1785
core:                  2250.00        Matches:         427
ercent Similarity:     92.64%         Conservative:    39
est Local Similarity:  84.89%         Mismatches:      33
iery Match:            86.04%         Indels:          4
3:                     8              Gaps:            2

3-10-006-852-2 (1-502) x PETGADX (1-1785)

            1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||      ||||||:::||||||||||||:::||||||||||||||||||
           72 ATGGTTCTATCAAAGACAGTGTCGCAGAGCGATGTGTCCATTCACTCCACGTTTGCTTCT 131

           21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
          132 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGATAATTCGATACCAAAAGAA 191

           41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          192 GCAGCATATCAGATCATAAATGATGAACTGATGTTAGATGGAAACCCAAGGCTGAACTTG 251

           61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
          252 GCTTCTTTTGTTACAACATGGATGGAACCAGAGTGTGATAAGTTGATGATGGACTCTATT 311

           81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          312 AACAAGAACTATGTTGATATGGATGAATATCCTGTTACCACTGAGCTTCAGAATCGATGT 371

          101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              |||||||||||||||||||||||||||||||||||||||:::   |||||||||||||||
          372 GTAAACATGATAGCTCATTTGTTTAATGCACCACTTGAAGATGGAGAAACTGCAGTTGGA 431

          121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          432 GTTGGAACTGTTGGATCCTCTGAAGCCATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 491

          141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
          492 TGGCAGAACAAAATGAAAGCCCAAGGCAAACCCTGTGACAAGCCCAACATTGTTACTGGT 551

          161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          552 GCAAATGTCCAGGTGTGCTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 611

          181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||||||||||||||||||||||||||::::::||||||:::|||||||||
          612 GTAAAGCTTAGTGAAGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAGATGGTGGAT 671

          201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
          672 GAAAACACCATTTGTGTAGCTGCTATCTTAGGTTCCACCCTCAATGGAGAATTTGAAGAC 731

          221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
          732 GTTAAGCGCTTGAATGATCTCTTGGTCGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 791
```

```
Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        792 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCGTTCATTTACCCAGAGCTTGAG 851

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 TGGGACTTTAGATTGCCATTAGTGAAGAGCATTAATGTAAGTGGTCACAAATATGGTCTT 911

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              ||||||||||||||||||||||||:::||||||||||||:::|||||||||:::||||||
Db        912 GTCTATGCTGGTATTGGTTGGGTCGTTTGGAGGAACAAGGATGATTTGCCTGATGAACTT 971

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTTTCTAAA 1031

Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db       1032 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTATGAGGGTTAC 1091

Qy        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              :::||||||||||||||||||:::||||||      ||||||||||||||||||||||||
Db       1092 AAGAATGTGATGGAGAATTGTCAAGAAAATGCATCGGTACTAAGAGAAGGGCTAGAAAAG 1151

Qy        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
              |||   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
Db       1152 ACAGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTACCTTTAGTAGCATTCTCTCTT 1211

Qy        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
              ||||||:::      |||   |||||||||||||||:::   |||||||||:::||||||
Db       1212 AAAGACAACAGGCAACACAACGAGTTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1271

Qy        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1272 ATTGTTCCTGCATATACTATGCCACCAAACGCACAACACATTACAGTTCTCAGAGTTGTG 1331

Qy        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db       1332 ATCAGAGAAGATTTCTCCCGTACGCTTGCAGAACGACTGGTAAGAGACATCGAAAAAGTC 1391

Qy        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              :::   |||||||||   ||||||:::||||||      |||:::::::::   :::|||
Db       1392 CTTCATGAACTTGACACACTCCCTGCACGTGTCAATGCTAAGCTCGCTGTGGCCGAGGAG 1451

Qy        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
              ::::::   :::|||         :::            ||||||:::|||   :::
Db       1452 CAGGCGGCTGCGAATGGCAGCGAGGTGCAT---------AAGAAAACAGATAGCGAAGTG 1502

Qy        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
              |||   ::::::||||||   |||||||||||||||   :::   :::|||||||||:::
Db       1503 CAGTTGGAGATGATAACTGCATGGAAGAAGTTTGTTGAAGAAAAGAAGAAGAAGACTAAT 1562

Qy        500 GlyIleCys 502
                 :::|||
Db       1563 CGAGTTTGT 1571

RESULT 7
PETGLUDECA
LOCUS       PETGLUDECA               1785 bp    mRNA    linear   PLN 26-JUL-1993
DEFINITION  Petunia hybrida glutamate decarboxylase (gad) mRNA, complete cds.
ACCESSION   L16977
VERSION     L16977.1  GI:309679
KEYWORDS    calmodulin binding protein; glutamate decarboxylase.
SOURCE      Petunia x hybrida
  ORGANISM  Petunia x hybrida
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Petunia.
REFERENCE   1  (bases 1 to 1785)
  AUTHORS   Baum,G., Chen,Y., Arazi,T., Takatsuji,H. and Fromm,H.
  TITLE     A plant glutamate decarboxylase containing a calmodulin-binding
            domain: cloning sequence and functional analysis
  JOURNAL   J. Biol. Chem. (1993) In press
```

```
OMMENT      Original source text: Petunia hybrida flower corolla cDNA to mRNA.
EATURES              Location/Qualifiers
    source           1. .1785
                     /organism="Petunia x hybrida"
                     /mol_type="mRNA"
                     /db_xref="taxon:4102"
                     /tissue_type="corolla"
                     /dev_stage="flower"
    gene             1. .1785
                     /gene="gad"
    CDS              72. .1574
                     /gene="gad"
                     /EC_number="4.1.1.15"
                     /codon_start=1
                     /product="glutamate decarboxylase"
                     /protein_id="AAA33710.1"
                     /db_xref="GI:309680"
                     /translation="MVLSKTVSQSDVSIHSTFASRYVRTSLPRFKMPDNSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLEDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSEGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLVEKNKETGWDTPIHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWVVWRNKDDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGYEGYKNVMENCQENASVLREGLEKTGRFNIISKEIGVPLVAFSLKDNRQHNEFEIS
                     ETLRRFGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVRDIEKVLHELDTLPAR
                     VNAKLAVAEEQAAANGSEVHKKTDSEVQLEMITAWKKFVEEKKKKTNRVC"
ASE COUNT      550 a    298 c    401 g    536 t
RIGIN

lignment Scores:
red. No.:              3.5e-181       Length:         1785
core:                  2250.00        Matches:        427
ercent Similarity:     92.64%         Conservative:   39
est Local Similarity:  84.89%         Mismatches:     33
uery Match:            86.04%         Indels:         4
B:                     8              Gaps:           2

S-10-006-852-2 (1-502) x PETGLUDECA (1-1785)

y           1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||      ||||||:::||||||||||||:::||||||||||||||||||
           72 ATGGTTCTATCAAAGACAGTGTCGCAGAGCGATGTGTCCATTCACTCCACGTTTGCTTCT 131

y          21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
          132 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGATAATTCGATACCAAAAGAA 191

y          41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          192 GCAGCATATCAGATCATAAATGATGAACTGATGTTAGATGGAAACCCAAGGCTGAACTTG 251

y          61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
          252 GCTTCTTTTGTTACAACATGGATGGAACCAGAGTGTGATAAGTTGATGATGGACTCTATT 311

y          81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          312 AACAAGAACTATGTTGATATGGATGAATATCCTGTTACCACTGAGCTTCAGAATCGATGT 371

y         101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              |||||||||||||||||||||||||||||||||||||||:::   |||||||||||||||
          372 GTAAACATGATAGCTCATTTGTTTAATGCACCACTTGAAGATGGAGAAACTGCAGTTGGA 431

y         121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          432 GTTGGAACTGTTGGATCCTCTGAAGCCATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 491

y         141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
          492 TGGCAGAACAAAATGAAAGCCCAAGGCAAACCCTGTGACAAGCCCAACATTGTTACTGGT 551

y         161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          552 GCAAATGTCCAGGTGTGCTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 611

Qy          181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
               ||||||||||||||||||||||||||||||||||||::::::||||||:::|||||||||
Db          612 GTAAAGCTTAGTGAAGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAGATGGTGGAT 671

Qy          201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db          672 GAAAACACCATTTGTGTAGCTGCTATCTTAGGTTCCACCCTCAATGGAGAATTTGAAGAC 731

Qy          221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
               ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          732 GTTAAGCGCTTGAATGATCTCTTGGTCGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 791

Qy          241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
               ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db          792 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCGTTCATTTACCCAGAGCTTGAG 851

Qy          261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          852 TGGGACTTTAGATTGCCATTAGTGAAGAGCATTAATGTAAGTGGTCACAAATATGGTCTT 911

Qy          281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
               ||||||||||||||||||||||||:::||||||||||||:::|||||||||:::||||||
Db          912 GTCTATGCTGGTATTGGTTGGGTCGTTTGGAGGAACAAGGATGATTTGCCTGATGAACTT 971

Qy          301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          972 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTTTCTAAA 1031

Qy          321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
               ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db         1032 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTATGAGGGTTAC 1091

Qy          341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
               :::||||||||||||||||||:::||||||      ||||||||||||||||||||||||
Db         1092 AAGAATGTGATGGAGAATTGTCAAGAAAATGCATCGGTACTAAGAGAAGGGCTAGAAAAG 1151

Qy          361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
               |||   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
Db         1152 ACAGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTACCTTTAGTAGCATTCTCTCTT 1211

Qy          381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
               ||||||:::      |||   ||||||||||||||||||:::   |||||||||:::||||||
Db         1212 AAAGACAACAGGCAACACAACGAGTTCGAGATTTCTGAAACTTTAAGGAGATTTGGTTGG 1271

Qy          401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1272 ATTGTTCCTGCATATACTATGCCACCAAACGCACAACACATTACAGTTCTCAGAGTTGTG 1331

Qy          421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
               ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db         1332 ATCAGAGAAGATTTCTCCCGTACGCTTGCAGAACGACTGGTAAGAGACATCGAAAAAGTC 1391

Qy          441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
               :::   |||||||||   ||||||:::||||||      |||:::::::::   :::|||
Db         1392 CTTCATGAACTTGACACACTCCCTGCACGTGTCAATGCTAAGCTCGCTGTGGCCGAGGAG 1451

Qy          461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
               ::::::   :::|||         :::            ||||||:::|||   :::
Db         1452 CAGGCGGCTGCGAATGGCAGCGAGGTGCAT---------AAGAAAACAGATAGCGAAGTG 1502

Qy          481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
               |||   ::::::||||||   |||||||||||||||   :::   :::|||||||||:::
Db         1503 CAGTTGGAGATGATAACTGCATGGAAGAAGTTTGTTGAAGAAAAGAAGAAGAAGACTAAT 1562

Qy          500 GlyIleCys 502
                  :::|||
Db         1563 CGAGTTTGT 1571

RESULT 8
AF020425
LOCUS       AF020425                  1705 bp    mRNA    linear   PLN 24-JUN-1998
```

```
EFINITION   Nicotiana tabacum glutamate decarboxylase isozyme 1 (NtGAD1) mRNA,
            complete cds.
CCESSION    AF020425
ERSION      AF020425.1  GI:3252855
EYWORDS     .
OURCE       Nicotiana tabacum (common tobacco)
 ORGANISM   Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
EFERENCE    1  (bases 1 to 1705)
 AUTHORS    Yun,S.J. and Oh,S.H.
 TITLE      Cloning and characterization of a tobacco cDNA encoding
            calcium/calmodulin-dependent glutamate decarboxylase
 JOURNAL    Unpublished
EFERENCE    2  (bases 1 to 1705)
 AUTHORS    Yun,S.J.
 TITLE      Direct Submission
 JOURNAL    Submitted (21-AUG-1997) Faculty of Biological Resources Science,
            Chonbuk National University, 664-14 1-ga Tokjin-dong, Chonju,
            Chonbuk 561-756, South Korea
EATURES              Location/Qualifiers
    source           1. .1705
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4097"
    gene             1. .1705
                     /gene="NtGAD1"
    CDS              71. .1561
                     /gene="NtGAD1"
                     /function="calmodulin binding protein"
                     /note="calcium-calmodulin-dependent enzyme"
                     /codon_start=1
                     /product="glutamate decarboxylase isozyme 1"
                     /protein_id="AAC24195.1"
                     /db_xref="GI:3252856"
                     /translation="MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECNKLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLGDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLIEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWAIWRNKEDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGFEGYKNVMENCQENARVLREGLEKSGRFNIISKEIGVPLVAFSLKDNSQHNEFEIS
                     ETLRRFGWIIPAYTMPPNAQHVTVLRVVIREDFSRTLAERLVIDIEKVLHELDTLPAR
                     VNAKLAVAEANGSGVHKKTDREVQLEITTAWKKFVADKKKKTNGVC"
ASE COUNT      523 a    301 c    385 g    496 t
RIGIN

lignment Scores:
red. No.:               2.32e-180      Length:         1705
core:                   2240.00        Matches:        428
ercent Similarity:      92.05%         Conservative:   35
est Local Similarity:   85.09%         Mismatches:     32
ery Match:              85.66%         Indels:         8
:                       8              Gaps:           3

-10-006-852-2 (1-502) x AF020425 (1-1705)

/          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
             ||||||||||||         ||||||||||||||||||:::|||||||||||||||||| 
)         71 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 130

/         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)        131 CGATATGTTCGTACTTCTCTTCCGAGGTTTAAGATGCCAGAGAATTCGATACCAAAGGAA 190

/         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
)        191 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGACTAAATTTA 250

/         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
             ||||||||||||||||||||||||||||||||||||:::||||||:::|||   ||||||
)        251 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAACAAACTGATGATGGATTCCATT 310

Qy        81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       311 AACAAGAATTACGTTGACATGGATGAATACCCTGTAACCACTGAACTTCAGAATCGATGT 370

Qy       101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
             ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
Db       371 GTAAACATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 430

Qy       121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       431 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 490

Qy       141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
             ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
Db       491 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTGTGACAAGCCCAATATTGTCACTGGT 550

Qy       161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       551 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 610

Qy       181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
             ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
Db       611 GTAAAGTTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 670

Qy       201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
             |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db       671 GAGAACACAATTTGTGTAGCTGCTATCTTGGGTTCCACACTCAATGGTGAATTTGAAGAT 730

Qy       221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
             ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
Db       731 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 790

Qy       241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       791 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 850

Qy       261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       851 TGGGACTTTAGATTGCCATTGGTGAAGAGTATAAACGTGAGTGGTCACAAATATGGTCTT 910

Qy       281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
             |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
Db       911 GTTTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTACCTGACGAACTT 970

Qy       301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       971 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1030

Qy       321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
             ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db      1031 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1090

Qy       341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
             :::||||||||||||||||||:::||||||      ||||||||||||||||||||||||
Db      1091 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1150

Qy       361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
             :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
Db      1151 AGTGGAAGATTCAACATAATATCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1210

Qy       381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
             ||||||:::|||   |||   |||||||||||||||:::   |||||||||:::||||||
Db      1211 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1270

Qy       401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
             |||:::|||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db      1271 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1330

Qy       421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1331 ATTAGAGAAGATTTCTCCCGTACACTCGCCGAGCGACTGGTAATAGACATTGAAAAAGTC 1390
```

```
y         441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
              :::   |||||||||   ||||||:::||||||      |||:::::::::
b        1391 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTG--------- 1441

y         461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                 :::|||:::|||         :::            ||||||:::|||   :::
b        1442 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1489

y         481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
              |||   :::|||   |||   |||||||||||||||||||||   :::|||||||||:::
ɔ        1490 CAGCTTGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAC 1549

y         500 GlyIleCys 502
              |||:::|||
ɔ        1550 GGAGTTTGT 1558

ESULT 9
F506366
OCUS       AF506366                1745 bp    mRNA    linear   PLN 02-MAY-2003
EFINITION  Nicotiana tabacum glutamate decarboxylase mRNA, complete cds.
CCESSION   AF506366
ERSION     AF506366.1  GI:21327028
EYWORDS    .
OURCE      Nicotiana tabacum (common tobacco)
 ORGANISM  Nicotiana tabacum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
EFERENCE   1  (bases 1 to 1745)
 AUTHORS   McLean,M.D., Yevtushenko,D., Deschene,A., Van Cauwenberghe,O.R.,
           Makhmoudova,A., Potter,J.W., Bown,A.W. and Shelp,B.J.
 TITLE     Overexpression of glutamate decarboxylase in transgenic tobacco
           confers resistance to the northern root-knot nemtaode
 JOURNAL   Mol. Breed. 11, 277-285 (2003)
EFERENCE   2  (bases 1 to 1745)
 AUTHORS   McLean,M.D., Yevtushenko,D. and Shelp,B.J.
 TITLE     Direct Submission
 JOURNAL   Submitted (25-APR-2002) Plant Agriculture, University of Guelph,
           Guelph, ON N1G 2W1, Canada
EATURES             Location/Qualifiers
    source          1. .1745
                    /organism="Nicotiana tabacum"
                    /mol_type="mRNA"
                    /cultivar="Samsun NN"
                    /db_xref="taxon:4097"
    CDS             31. .1521
                    /note="GAD20.1"
                    /codon_start=1
                    /product="glutamate decarboxylase"
                    /protein_id="AAM48129.1"
                    /db_xref="GI:21327029"
                    /translation="MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                    IINDELMLDGNPRLNLASFVTTWMEPECNTLMMDSINKNYVDMDEYPVTTELQNRCVN
                    MIAHLFNAPLGDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPFDKPNIVTG
                    ANVQVCWEKFARYFEVELKEVKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                    EDVKRLNDLLIEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                    KYGLVYAGIGWAIWRNKEDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                    LGFEGYKNVMENCQENARVLREGIEKSGRFNIISKEIGVPLVAFSLKDNSQHNEFEIS
                    ETLRRFGWIVPAYTMPPNAQHVTVLRVVIREDFSRTLAERLVIDIEKVLHELDTLPAR
                    VNAKLAVAEANGSGVHKKTDREVQLEITTAWKKFVADKKKKTNGVC"
ASE COUNT      555 a    290 c    389 g    511 t
RIGIN

lignment Scores:
red. No.:              9.37e-180     Length:        1745
core:                  2233.00       Matches:       427
ercent Similarity:     91.85%        Conservative:  35
est Local Similarity:  84.89%        Mismatches:    33
uery Match:            85.39%        Indels:        8
3:                     8             Gaps:          3

3-10-006-852-2 (1-502) x AF506366 (1-1745)
```

```
Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||         ||||||||||||||||||:::||||||||||||||||||
Db         31 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 90

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAGATGCCAGAGAATTCAATACCAAAGGAA 150

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 GCAGCATATCAGATTATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 210

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              ||||||||||||||||||||||||||||||||||||:::   |||:::|||   ||||||
Db        211 GCATCTTTCGTTACAACATGGATGGAGCCAGAATGTAATACGTTAATGATGGATTCCATT 270

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 AACAAGAACTACGTTGACATGGATGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 330

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
Db        331 GTAAATATGATAGCTCATTTGTTTAATGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 390

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 GTTGGAACTGTTGGATCCTCTGAAGCTATTATGCTTGCTGGATTAGCCTTTAAAAGAAAA 450

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
Db        451 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTTTGATAAGCCCAATATCGTCACCGGT 510

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 GCTAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 570

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
Db        571 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 630

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        631 GAGAATACCATTTGTGTTGCTGCTATCTTAGGTTCAACACTCAATGGTGAATTTGAAGAT 690

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
Db        691 GTTAAGCGTTTGAATGACCTTTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 750

Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        751 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 810

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        811 TGGGACTTTAGATTGCCATTGGTGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 870

Qy        281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
              |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
Db        871 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 930

Qy        301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        931 ATTTTCCACATCAATTACCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 990

Qy        321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db        991 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1050

Qy        341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
              :::||||||||||||||||||:::||||||      |||||||||||||||:::||||||
Db       1051 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTATTAAGAGAAGGAATTGAAAAA 1110

Qy        361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
```

```
            :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
       1111 AGTGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTTCCCTTAGTAGCATTTTCTCTT 1170

        381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
            ||||||:::|||   |||   |||||||||||||||:::   |||||||||:::||||||
       1171 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1230

        401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
            |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
       1231 ATTGTTCCTGCATATACTATGCCACCAAATGCTCAACATGTTACAGTTCTCAGAGTTGTC 1290

        421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       1291 ATTAGAGAAGATTTCTCCCGCACACTAGCGGAGCGACTGGTAATAGACATTGAAAAAGTC 1350

        441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
            :::   |||||||||   ||||||:::||||||      |||:::::::::
       1351 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTG--------- 1401

        461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
               :::|||:::|||         :::            ||||||:::|||   :::
       1402 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1449

        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
            |||   :::|||   |||   |||||||||||||||||||||   :::|||||||||:::
       1450 CAGCTAGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAT 1509

        500 GlyIleCys 502
            |||:::|||
       1510 GGAGTTTGT 1518

SULT 10
U54774
CUS        NTU54774                1929 bp    mRNA    linear   PLN 14-JAN-1997
FINITION   Nicotiana tabacum glutamate decarboxylase (NtGAD1) mRNA, complete
           cds.
CESSION    U54774
RSION      U54774.1  GI:1777920
YWORDS     .
URCE       Nicotiana tabacum (common tobacco)
 ORGANISM  Nicotiana tabacum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
FERENCE    1  (bases 1 to 1929)
 AUTHORS   Dharmasiri,M.A.N., Lu,Y.T. and Harrington,H.M.
 TITLE     Cloning and sequencing of a tobacco cDNA encoding glutamate
           decarboxylase
 JOURNAL   Unpublished
FERENCE    2  (bases 1 to 1929)
 AUTHORS   Dharmasiri,M.A.N., Lu,Y.T. and Harrington,H.M.
 TITLE     Direct Submission
 JOURNAL   Submitted (11-APR-1996) M.A. Nihal Dharmasiri, Plant Molecular
           Physiology, University of Hawaii, 3190 Maile Way, Honolulu, HI
           96822, USA
ATURES              Location/Qualifiers
   source           1..1929
                    /organism="Nicotiana tabacum"
                    /mol_type="mRNA"
                    /strain="Xanthi"
                    /db_xref="taxon:4097"
   gene             1..1929
                    /gene="NtGAD1"
   CDS              66..1556
                    /gene="NtGAD1"
                    /note="calmodulin regulated enzyme; calmodulin-binding
                    protein"
                    /codon_start=1
                    /product="glutamate decarboxylase"
                    /protein_id="AAB40608.1"
                    /db_xref="GI:1777921"
                    /translation="MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                    IINDELMLDGNPRLNLASFVTTWMEPECNKLMMDSINKNYVDMDEYPVTTELQNRCVN
```

```
                    MIAHLFNAPLGDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                    ANVQVCWEKFARYFEVELKEVKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                    EDVKRLNDLLIEKNKETGWDTPIHVDAASGEFIAPFLYPELEWDFRLPLVKSINVSGH
                    KYGLVYAGIGWAIWRNKEDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                    LGFEGYKNVMENCQENARVLREGLEKSGRFNIISKEIGVPLVAFSLKDNSQHNEFEIS
                    ETLRRFGWIIPAYTMPPNAQHVTVLRVVIREDFSRTLAERLVIDIEKVLHELDTLPAR
                    VNAKLAVAEANGSGVHKKTDREVQLEITTAWKKFVADKKKKTNGVC"
BASE COUNT      607 a    323 c    416 g    583 t
ORIGIN

Alignment Scores:
Pred. No.:               1.29e-179      Length:          1929
Score:                   2232.00        Matches:         427
Percent Similarity:      91.85%         Conservative:    35
Best Local Similarity:   84.89%         Mismatches:      33
Query Match:             85.35%         Indels:          8
DB:                      8              Gaps:            3

US-10-006-852-2 (1-502) x NTU54774 (1-1929)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||         ||||||||||||||||||:::||||||||||||||||||
Db         66 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 125

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 CGATATGTTCGTACTTCTCTTCCGAGGTTTAAGATGCCAGAGAATTCGATACCAAAGGAA 185

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGACTAAATTTA 245

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              ||||||||||||||||||||||||||||||||||||:::||||||:::|||   ||||||
Db        246 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAACAAACTGATGATGGATTCCATT 305

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 AACAAGAATTACGTTGACATGGATGAATACCCTGTAACCACTGAACTTCAGAATCGATGT 365

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
Db        366 GTAAACATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 425

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCTTTCAAGAGAAAA 485

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
Db        486 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTGTGACAAGCCCAATATTGTCACTGGT 545

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGCTAAAGGAA 605

Qy        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
              ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
Db        606 GTAAAGTTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 665

Qy        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        666 GAGAACACAATTTGTGTAGCTGCTATCTTGGGTTCCACACTCAATGGTGAATTTGAAGAT 725

Qy        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
              ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
Db        726 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 785

Qy        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        786 ATTCATGTGGATGCAGCAAGTGGTGAATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 845

Qy        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 846 TGGGACTTTAGATTGCCATTGGTGAAGAGTATTAACGTGAGTGGTCACAAATATGGTCTT 905

 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
     |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
 906 GTTTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTACCTGACGAACTT 965

 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 966 ATCTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1025

 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
     ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
1026 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1085

 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
     :::||||||||||||||||||:::||||||      ||||||||||||||||||||||||
1086 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1145

 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
     :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
1146 AGTGGAAGATTCAACATAATATCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1205

 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
     ||||||:::|||   |||   ||||||||||||:::   |||||||||:::||||||
1206 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1265

 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
     |||:::|||||||||||||||||||||||||||||||||:::||||||||||||||||||
1266 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1325

 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1326 ATTAGAGAAGATTTCTCCCGTACACTCGCCGAGCGACTGGTAATAGACATTGAAAAAGTC 1385

 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
     :::   |||||||||   ||||||:::||||||      |||:::::::::
1386 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTG--------- 1436

 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
        :::|||:::|||         :::            ||||||:::|||   :::
1437 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1484

 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
     |||   :::|||   |||   |||||||||||||||||||||   :::|||||||||:::
1485 CAGCTTGAGATTACTACTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAC 1544

 500 GlyIleCys 502
     |||:::|||
1545 GGAGTTTGT 1553

SULT 11
352732
CUS        AF352732                1672 bp    mRNA    linear   PLN 02-MAY-2003
FINITION   Nicotiana tabacum glutamate decarboxylase isozyme 1 mRNA, complete
           cds.
CESSION    AF352732
RSION      AF352732.1  GI:13310812
YWORDS     .
URCE       Nicotiana tabacum (common tobacco)
 ORGANISM  Nicotiana tabacum
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
CFERENCE   1  (bases 1 to 1672)
 AUTHORS   Yevtushenko,D., McLean,M.D., Peiris,S.E., Van Cauwenberghe,O.R. and
           Shelp,B.J.
 TITLE     Two isoforms of tobacco glutamate decarboxylase are regulated by
           calcium/calmodulin and differ in organ distribution
 JOURNAL   Unpublished
CFERENCE   2  (bases 1 to 1672)
 AUTHORS   Yevtushenko,D., McLean,M.D. and Shelp,B.J.
 TITLE     Direct Submission
```

```
  JOURNAL   Submitted (23-FEB-2001) Department of Plant Agriculture, University
            of Guelph, Division of Biotechnology, Guelph, Ontario N1G 2W1,
            Canada
FEATURES             Location/Qualifiers
     source          1. .1672
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /cultivar="Samsun NN"
                     /db_xref="taxon:4097"
     CDS             4. .1494
                     /EC_number="4.1.1.15"
                     /function="converts glutamate to gamma-aminobutyric acid"
                     /note="GAD; GAD3; NtGAD3; calcium/calmodulin-dependent
                     enzyme"
                     /codon_start=1
                     /product="glutamate decarboxylase isozyme 1"
                     /protein_id="AAK18620.1"
                     /db_xref="GI:13310813"
                     /translation="MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECNKLMMDSINKNYVDMGEYPVTTELQNRCVN
                     MIAHLFNAPLGDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLIEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
                     KYGLVYAGIGWAIWRNKEDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGFEGYKNVMENCQENARVLREGLEKSGRFNIISKEIGVPLVAFSLKDNSQHNEFEIS
                     ETLRRFGWIIPAYTMPPNAQHVTVLRVVIREDFSRTLAERLVIDIEKVLHELDTLPAR
                     VNAKLAVAEANGSGVHKKTDREVQLEITAAWKKFVADKKKKTNGVC"
BASE COUNT     511 a    280 c    392 g    483 t      6 others
ORIGIN

Alignment Scores:
Pred. No.:              2.36e-179       Length:         1672
Score:                  2228.00         Matches:        426
Percent Similarity:     91.65%          Conservative:   35
Best Local Similarity:  84.69%          Mismatches:     34
Query Match:            85.20%          Indels:         8
DB:                     8               Gaps:           3

US-10-006-852-2 (1-502) x AF352732 (1-1672)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||         ||||||||||||||||||:::||||||||||||||||||
Db          4 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCATCCACTCCACTTTCGCTTCC 63

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAGATGCCAGAGAATTCTATACCAAAAGAA 123

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 GCAGCATATCAAATCATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 183

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              ||||||||||||||||||||||||||||||||||||:::||||||:::|||   ||||||
Db        184 GCATCTTTTGTGACAACATGGATGGAACCAGAGTGTAATAAGTTAATGATGGATTCCATT 243

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        244 AACAAGAACTACGTTGACATGGGTGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 303

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
              ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
Db        304 GTAAATATGATAGCTCATTTGTTTAACGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 363

Qy        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 GTTGGAACTGTTGGATCCTCTGAGGCTATTATGCTTGCTGGATTAGCCTTCAAGAGAAAA 423

Qy        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
              ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
Db        424 TGGCAAAATAAAATGAAAGCCCAAGGGAAGCCCTGCGACAAGCCCAATATTGTCACTGGT 483

Qy        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 484 GCCAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 543

 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
     ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
 544 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 603

 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
     |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
 604 GAGAATACAATTTGTGTAGCTGCTATCTTGGGTTCCACTCTCAATGGTGAATTTGAAGAT 663

 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
     ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
 664 GTTAAGCGCTTGAATGACCTCTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 723

 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 724 ATTCATGTGGATGCAGCAAGTGGTGGATTCATTGCACCATTCCTTTATCCTGAGCTTGAA 783

 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 784 TGGGATTTTAGATTACCATTGGTGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 843

 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
     |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
 844 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 903

 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 904 ATTTTCCACATTAATTATCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 963

 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
     ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
 964 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1023

 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
     :::||||||||||||||||||:::||||||      ||||||||||||||||||||||||
1024 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTACTAAGAGAAGGACTTGAAAAA 1083

 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
     :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
1084 AGTGGAAGATTCAATATAATCTCCAAAGAAATTGGAGTTCCATTAGTAGCTTTCTCTCTT 1143

 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
     ||||||:::|||   |||   |||||||||||||||:::   |||||||||:::||||||
1144 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1203

 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
     |||:::|||||||||||||||||||||||||||||||||:::||||||||||||||||||
1204 ATTATTCCTGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTT 1263

 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1264 ATTAGAGAAGATTTCTCCCGTACACTCGCGGAGCGACTGGTGATAGACATTGAAAAAGTC 1323

 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
     :::   |||||||||   ||||||:::||||||      |||:::::::::
1324 CTCCACGAGCTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTCGCCGTG--------- 1374

 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
        :::|||:::|||         :::            ||||||:::|||   :::
1375 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1422

 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
     |||   :::|||         |||||||||||||||||||||   :::|||||||||:::
1423 CAGCTGGAGATTACTGCTGCATGGAAGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAT 1482

 500 GlyIleCys 502
     |||:::|||
1483 GGAGTTTGT 1491

ESULT 12
X654603
```

```
LOCUS       AX654603                1479 bp    DNA     linear   PAT 22-MAR-2003
DEFINITION  Sequence 4473 from Patent WO03000898.
ACCESSION   AX654603
VERSION     AX654603.1  GI:29157417
KEYWORDS    .
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
            Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
  TITLE     Plant genes involved in defense against pathogens
  JOURNAL   Patent: WO 03000898-A 4473 03-JAN-2003;
            Syngenta Participations AG (CH)
FEATURES             Location/Qualifiers
     source          1..1479
                     /organism="Oryza sativa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4530"
BASE COUNT      298 a    478 c    472 g    231 t
ORIGIN

Alignment Scores:
Pred. No.:              4.77e-177     Length:          1479
Score:                  2200.00       Matches:         420
Percent Similarity:     90.64%        Conservative:    35
Best Local Similarity:  83.67%        Mismatches:      37
Query Match:            84.13%        Indels:          10
DB:                     6             Gaps:            2

US-10-006-852-2 (1-502) x AX654603 (1-1479)

Qy         1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
             ||||||||||||   ||||||||||||||||||:::||||||||||||||||||||||||
Db         1 ATGGTGCTCTCCAAGGCCGTCTCCGAGAGTGACATGTCCGTGCACTCCACCTTCGCCTCC 60

Qy        21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
             ||||||||||||   ||||||||||||::::::|||||||||||||||||||||||||||
Db        61 CGCTACGTCCGCGCCTCCCTCCCAAGGTACCGGATGCCGGAGAACTCGATCCCGAAGGAG 120

Qy        41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GCGGCGTACCAGATCATCAACGACGAGCTGATGCTGGACGGCAACCCGCGGCTGAACCTG 180

Qy        61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
             |||||||||||||||||||||||||||||||||||||||||||||||||||::::::|||
Db       181 GCGTCGTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATCATGGCCGCCATC 240

Qy        81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 AACAAGAACTACGTCGACATGGACGAGTACCCCGTCACCACCGAGCTCCAGAACCGGTGC 300

Qy       101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
             ||||||||||||||||||||||||:::|||||||||   |||   |||||||||||||||
Db       301 GTGAACATGATCGCGCACCTGTTCCACGCGCCGCTCGGGGAGGACGAGACGGCGGTGGGC 360

Qy       121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       361 GTGGGCACGGTGGGTTCGTCGGAGGCCATCATGCTGGCCGGGCTGGCCTTCAAGCGGCGG 420

Qy       141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
             |||||||||||||||||||||||||||||||||   |||||||||||||||:::||||||
Db       421 TGGCAGAACAAGCGCAAGGCCGAGGGGAAGCCGTTCGACAAGCCCAACATCATCACCGGC 480

Qy       161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GCCAACGTGCAGGTGTGCTGGGAGAAGTTCGCCCGCTACTTCGAGGTGGAGCTCAAGGAG 540

Qy       181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
             |||||||||   :::|||||||||||||||||||||::::::|||||||||||||||:::
Db       541 GTGAAGCTCCGCGACGGCTACTACGTCATGGACCCCGAGAAGGCCGTCGACATGGTCAAC 600
```

```
 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
     |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
 601 GAGAACACCATCTGCGTCGCCGCCATCCTCGGCTCCACCCTCAACGGCGAGTTCGAGGAC 660

 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
     ||||||||||||||||||||||||   :::|||||||||||||||||||||:::||||||
 661 GTCAAGCTACTCAACGACCTCCTCGACAAGAAGAACAAGGAGACTGGGTGGGAGACGCCG 720

 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 721 ATCCACGTGGACGCGGCGAGCGGCGGGTTCATCGCGCCGTTCCTGTACCCGGAGCTGGAG 780

 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
     ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
 781 TGGGACTTCCGGCTGCCGTGGGTGAAGAGCATCAACGTGAGCGGTCACAAGTACGGGCTC 840

 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
     |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
 841 GTCTACGCCGGCATCGGCTGGTGCATCTGGCGCAACAAGGAGGACCTGCCCGAGGAGCTC 900

 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
     ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
 901 ATCTTCCACATCAACTACCTCGGCACCGACCAGCCAACCTTCACCCTCAACTTCTCCAAG 960

 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
     ||||||||||||||||||||||||||||||||||||||||||   |||   |||||||||
 961 GGCTCCAGCCAGGTCATCGCCCAGTACTACCAGCTCATCCGCCACGGCTTCGAGGGGTAC 1020

 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
     ||||||:::||||||||||||   ||||||   :::||||||:::|||||||||   |||
1021 AGGAACATCATGGAGAACTGCCACGAGAACGCGATGGTGCTGAAGGAAGGGCTGGTGAAG 1080

 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
     |||   ||||||:::|||||||||||||||||||||||||||||||||||||||||||||
1081 ACGGGGAGGTTCGACATCGTGTCCAAGGACGAAGGGGTGCCGCTGGTGGCGTTCTCGCTC 1140

 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
     ||||||   |||   |||   ||||||||||||||||||||||||||||||:::||||||
1141 AAGGACCGGAGCCGGCACGACGAGTTCGAGATCTCCGACATGCTGCGCCGCTTCGGCTGG 1200

 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
     |||||||||||||||||||||||||||:::|||||||||:::||||||||||||||||||
1201 ATCGTGCCGGCGTACACCATGCCGCCCGACGCCCAGCACGTCACGGTGCTCCGCGTGGTC 1260

 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
     |||||||||:::||||||||||||||||||||||||||||||:::|||||||||||||||
1261 ATCCGGGAGGAGTTCAGCCGCACCCTCGCCGAGCGCCTCGTCCTCGACATCGAGAAGGTG 1320

 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
     |||   :::||||||   ||||||||||||::::::      :::
1321 ATGTACCAGCTCGACGCGCTCCCCTCCAGGCTCATGCCCCCCGTGCCGCCGGCGCCG--- 1377

 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                             |||:::|||      |||||||||:::::::::
1378 ------------------------CTGCTGGTGGTCGCCAAGAAGTCGGAGCTCGAGACG 1413

 481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
     ||||||   :::         |||||||||||||||         |||:::|||:::|||
1414 CAGCGGTCGGTGACGGAGGCGTGGAAGAAGTTCGTGCTC---GCCAAGAGGACCAACGGC 1470

 501 IleCys 502
     :::|||
1471 GTCTGC 1476

ESULT 13
Y187941
OCUS       AY187941                 1502 bp    mRNA    linear   PLN 10-MAR-2003
EFINITION  Oryza sativa (japonica cultivar-group) glutamate decarboxylase
           (GAD3) mRNA, complete cds.
CCESSION   AY187941
ERSION     AY187941.1  GI:28911952
EYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1502)
  AUTHORS   Liu,L.L., Zhai,H.Q. and Wan,J.M.
  TITLE     New rice GAD gene GAD3
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1502)
  AUTHORS   Liu,L.L., Zhai,H.Q. and Wan,J.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-NOV-2002) State Key Laboratory of Crop Genetics
            Germplasm Enhancement, Jiangsu Plant Gene Engineering Research
            Center, Weigang, Nanjing, Jiangsu 210095, P. R. China
FEATURES             Location/Qualifiers
     source          1. .1502
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Haiminori"
                     /db_xref="taxon:39947"
     gene            1. .1502
                     /gene="GAD3"
     CDS             21. .1499
                     /gene="GAD3"
                     /EC_number="4.1.1.15"
                     /codon_start=1
                     /product="glutamate decarboxylase"
                     /protein_id="AAO59316.1"
                     /db_xref="GI:28911953"
                     /translation="MVLSKAVSESDMSVHSTFASRYVRASLPRYRMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLIMAAINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFHAPLGEDETAVGVGTVGSSEAIMLAGLAFKRRWQNKRKAEGKPFDKPNIITG
                     ANVQVCWEKFARYFEVELKEVKLRDGYYVMDPEKAVDMVNENTICVAAILGSTLNGEF
                     EDVKLLNDLLDKKNKETGWETPIHVDAASGGFIAPFLYPELEWDFRLPWVKSINVSGH
                     KYGLVYAGIGWCIWRNKEDLPEELIFHINYLGTDQPTFTLNFSKGSSQVIAQYYQLIR
                     HGFEGYRNIMENCHENAMVLKEGLVKTGRFDIVSKDEGVPLVAFSLKDRSRHDEFEIS
                     DMLRRFGWIVPAYTMPPDAQHVTVLRVVIREEFSRTLAERLVLDIEKVMYQLDALPSR
                     LMPPVPPAPLLVVAKKSELETQRSVTEAWKKFVLAKRTNGVC"
BASE COUNT      303 a    483 c    479 g    237 t
ORIGIN

Alignment Scores:
Pred. No.:              4.86e-177       Length:          1502
Score:                  2200.00         Matches:         420
Percent Similarity:     90.64%          Conservative:    35
Best Local Similarity:  83.67%          Mismatches:      37
Query Match:            84.13%          Indels:          10
DB:                     8               Gaps:            2

US-10-006-852-2 (1-502) x AY187941 (1-1502)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
              ||||||||||||   ||||||||||||||||||:::||||||||||||||||||||||||
Db         21 ATGGTGCTCTCCAAGGCCGTCTCCGAGAGTGACATGTCCGTGCACTCCACCTTCGCCTCC 80

Qy         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
              ||||||||||||   ||||||||||||::::::|||||||||||||||||||||||||||
Db         81 CGCTACGTCCGCGCCTCCCTCCCAAGGTACCGGATGCCGGAGAACTCGATCCCGAAGGAG 140

Qy         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 GCGGCGTACCAGATCATCAACGACGAGCTGATGCTGGACGGCAACCCGCGGCTGAACCTG 200

Qy         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
              |||||||||||||||||||||||||||||||||||||||||||||||||||::::::|||
Db        201 GCGTCGTTCGTCACCACGTGGATGGAGCCCGAGTGCGACAAGCTCATCATGGCCGCCATC 260

Qy         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 AACAAGAACTACGTCGACATGGACGAGTACCCCGTCACCACCGAGCTCCAGAACCGGTGC 320

Qy        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
```

```
     ||||||||||||||||||||||||:::|||||||||   |||   |||||||||||||||
 321 GTGAACATGATCGCGCACCTGTTCCACGCGCCGCTCGGGGAGGACGAGACGGCGGTGGGC 380

 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
     |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
 381 GTGGGCACGGTGGGTTCGTCGGAGGCCATCATGCTGGCCGGGCTGGCCTTCAAGCGGCGG 440

 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
     |||||||||||||||||||||||||||||||||   |||||||||||||||:::||||||
 441 TGGCAGAACAAGCGCAAGGCCGAGGGGAAGCCGTTCGACAAGCCCAACATCATCACCGGC 500

 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 501 GCCAACGTGCAGGTGTGCTGGGAGAAGTTCGCCCGCTACTTCGAGGTGGAGCTCAAGGAG 560

 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
     |||||||||   :::|||||||||||||||||||||::::::|||||||||||||||:::
 561 GTGAAGCTCCGCGACGGCTACTACGTCATGGACCCCGAGAAGGCCGTCGACATGGTCAAC 620

 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
     |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
 621 GAGAACACCATCTGCGTCGCCGCCATCCTCGGCTCCACCCTCAACGGCGAGTTCGAGGAC 680

 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
     ||||||||||||||||||||||||   :::|||||||||||||||||||||:::||||||
 681 GTCAAGCTACTCAACGACCTCCTCGACAAGAAGAACAAGGAGACTGGGTGGAGACGCCG 740

 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 741 ATCCACGTGGACGCGGCGAGCGGCGGGTTCATCGCGCCGTTCCTGTACCCGGAGCTGGAG 800

 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
     ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
 801 TGGGACTTCCGGCTGCCGTGGGTGAAGAGCATCAACGTGAGCGGTCACAAGTACGGGCTC 860

 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
     |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
 861 GTCTACGCCGGCATCGGCTGGTGCATCTGGCGCAACAAGGAGGACCTGCCCGAGGAGCTC 920

 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
     ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
 921 ATCTTCCACATCAACTACCTCGGCACCGACCAGCCAACCTTCACCCTCAACTTCTCCAAG 980

 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
     ||||||||||||||||||||||||||||||||||||||||||   |||   |||||||||
 981 GGCTCCAGCCAGGTCATCGCCCAGTACTACCAGCTCATCCGCCACGGCTTCGAGGGGTAC 1040

 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
     ||||||:::||||||||||||   ||||||   :::||||||:::|||||||||   |||
1041 AGGAACATCATGGAGAACTGCCACGAGAACGCGATGGTGCTGAAGGAAGGGCTGGTGAAG 1100

 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
     |||   ||||||:::|||||||||||||||||||||||||||||||||||||||||||||
1101 ACGGGGAGGTTCGACATCGTGTCCAAGGACGAAGGGGTGCCGCTGGTGGCGTTCTCGCTC 1160

 381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
     ||||||   |||   |||   ||||||||||||||||||||||||||||||:::||||||
1161 AAGGACCGGAGCCGGCACGACGAGTTCGAGATCTCCGACATGCTGCGCCGCTTCGGCTGG 1220

 401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
     |||||||||||||||||||||||||||:::|||||||||:::||||||||||||||||||
1221 ATCGTGCCGGCGTACACCATGCCGCCCGACGCCCAGCACGTCACGGTGCTCCGCGTGGTC 1280

 421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
     |||||||||:::||||||||||||||||||||||||||||||:::|||||||||||||||
1281 ATCCGGGAGGAGTTCAGCCGCACCCTCGCCGAGCGCCTCGTCCTCGACATCGAGAAGGTG 1340

 441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
     |||   :::||||||   ||||||||||||::::::      :::
1341 ATGTACCAGCTCGACGCGCTCCCCTCCAGGCTCATGCCCCCCGTGCCGCCGGCGCCG--- 1397

 461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                             |||:::|||      |||||||||:::::::::
```

```
Db       1398 ------------------------CTGCTGGTGGTCGCCAAGAAGTCGGAGCTCGAGACG 1433

Qy        481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLysLysThrSerGly 500
              ||||||   :::         |||||||||||||||         |||:::|||:::|||
Db       1434 CAGCGGTCGGTGACGGAGGCGTGGAAGAAGTTCGTGCTC---GCCAAGAGGACCAACGGC 1490

Qy        501 IleCys 502
              :::|||
Db       1491 GTCTGC 1496

RESULT 14
AF020424
LOCUS       AF020424                1771 bp    mRNA    linear   PLN 24-JUN-1998
DEFINITION  Nicotiana tabacum glutamate decarboxylase isozyme 2 (NtGAD2) mRNA,
            complete cds.
ACCESSION   AF020424
VERSION     AF020424.1  GI:3252853
KEYWORDS    .
SOURCE      Nicotiana tabacum (common tobacco)
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1771)
  AUTHORS   Yun,S.J. and Oh,S.H.
  TITLE     Cloning and characterization of a tobacco cDNA encoding
            calcium/calmodulin-dependent glutamate decarboxylase
  JOURNAL   Mol. Cells 8 (2), 125-129 (1998)
  MEDLINE   98302498
   PUBMED   9638642
REFERENCE   2  (bases 1 to 1771)
  AUTHORS   Yun,S.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-AUG-1997) Faculty of Biological Resources Science,
            Chonbuk National University, 664-14 1-ga Tokjin-dong, Chonju,
            Chonbuk 561-756, South Korea
FEATURES             Location/Qualifiers
     source          1. .1771
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4097"
     gene            1. .1771
                     /gene="NtGAD2"
     CDS             67. .1557
                     /gene="NtGAD2"
                     /note="calcium-calmodulin-dependent enzyme"
                     /codon_start=1
                     /product="glutamate decarboxylase isozyme 2"
                     /protein_id="AAC39483.1"
                     /db_xref="GI:3252854"
                     /translation="MVLSKTASESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQ
                     IINDELMLDGNPRLNLASFVTTWMEPECNTLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIAHLFNAPLGDGETAVGVGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPFDKPNIVTG
                     ANVQVCWEKFARYFEVELKEVKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEF
                     EDVKRLNDLLIEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRLPLEKSINVSGH
                     KYGLVYAGIGWAIWRNKEDLPDELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGFEGYKNVMENCQENARVLREGIEKSGRFNIISKEIGVPLVAFSLKDNSQHNEFEIS
                     ETLRRFGWIVLAYTMPPNAQHVTVLRVVIREDFSRTLAERLVIDIEKVFHGVDTLPAR
                     VNAKLAVAEANGSGVHKKTDREVQLEITTAWLKFVADKKKKTNGVC"
BASE COUNT      548 a    295 c    393 g    535 t
ORIGIN

Alignment Scores:
Pred. No.:              7.24e-177      Length:         1771
Score:                  2199.00        Matches:        423
Percent Similarity:     90.85%         Conservative:   34
Best Local Similarity:  84.10%         Mismatches:     38
Query Match:            84.09%         Indels:         8
DB:                     8              Gaps:           3

US-10-006-852-2 (1-502) x AF020424 (1-1771)

Qy          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
```

```
     ||||||||||||         |||||||||||||||||||||||||||||||||||||||
  67 ATGGTTCTGTCCAAGACAGCGTCGGAAAGTGACGTCTCCGTTCACTCCACTTTCGCCTCC 126

  21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 127 CGATATGTTCGAACTTCTCTTCCCAGGTTTAAAATGCCAGAGAATTCAATACCAAAGGAA 186

  41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 187 GCAGCATATCAGATTATAAATGATGAGCTTATGTTAGATGGAAATCCAAGGCTAAATTTA 246

  61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
     ||||||||||||||||||||||||||||||||||||:::   |||:::|||   ||||||
 247 GCATCTTTCGTTACAACATGGATGGAGCCAGAATGTAATACGTTAATGATGGATTCCATT 306

  81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 307 AACAAGAACTACGTTGACATGGATGAATACCCTGTAACCACTGAGCTTCAGAATCGATGT 366

 101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
     ||||||||||||||||||||||||||||||||||||   :::   |||||||||||||||
 367 GTAAATATGATAGCTCATTTGTTTAATGCACCACTTGGAGATGGAGAGACTGCAGTTGGA 426

 121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 427 GTTGGAACTGTTGGATCCTCTGAAGCTATTATGCTTGCTGGATTAGCCTTTAAGAGAAAA 486

 141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
     ||||||||||||   ||||||:::|||||||||   ||||||||||||||||||||||||
 487 TGGCAAAATAAAATGAAAGCCCAAGGCAAGCCCTTTGATAAGCCCAATATTGTCACCGGT 546

 161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 547 GCTAATGTCCAGGTGTGTTGGGAGAAATTTGCAAGGTATTTTGAAGTGGAGTTGAAAGAA 606

 181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
     ||||||||||||:::|||||||||||||||||||||::::::||||||:::|||||||||
 607 GTAAAATTGAGTGATGGATACTATGTGATGGACCCTGAGAAAGCTGTGGAAATGGTGGAT 666

 201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
     |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
 667 GAGAATACCATTTGTGTTGCTGCTATCTTAGGTTCAACACTCAATGGTGAATTTGAAGAT 726

 221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
     ||||||   |||||||||||||||:::|||||||||||||||||||||||||||||||||
 727 GTTAAGCGTTTGAATGACCTTTTGATTGAGAAGAACAAAGAAACCGGGTGGGACACTCCA 786

 241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 787 ATTCATGTGGATGCAGCAAGTGGTGGATTTATTGCACCATTCCTTTATCCAGAGCTTGAA 846

 261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
     |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
 847 TGGGACTTTAGATTGCCATTGGAGAAGAGTATTAATGTGAGTGGTCACAAATATGGTCTT 906

 281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
     |||||||||||||||||||||   |||||||||||||||||||||||||||:::||||||
 907 GTCTATGCTGGTATTGGTTGGGCCATTTGGAGGAATAAGGAAGACTTGCCTGATGAACTT 966

 301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 967 ATTTTCCACATCAATTACCTTGGTGCTGATCAACCTACTTTCACTCTCAACTTCTCTAAA 1026

 321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
     ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
1027 GGTTCTAGCCAAGTAATTGCTCAATATTACCAACTTATTCGCTTGGGTTTTGAGGGTTAC 1086

 341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
     :::||||||||||||||||||:::||||||      |||||||||||||||:::||||||
1087 AAGAATGTTATGGAGAATTGTCAAGAAAATGCAAGGGTATTAAGAGAAGGAATTGAAAAA 1146

 361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
     :::   ||||||||||||:::||||||:::   |||||||||||||||||||||||||||
```

```
Db  1147 AGTGGAAGATTCAACATAATCTCCAAAGAAATTGGAGTTCCCTTAGTAGCATTTTCTCTT 1206

Qy   381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
         ||||||:::|||   |||   |||||||||||||||:::   |||||||||:::||||||
Db  1207 AAAGACAACAGTCAACACAATGAGTTCGAAATTTCTGAAACTCTTAGAAGATTTGGATGG 1266

Qy   401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
         ||||||   ||||||||||||||||||||||||||||||:::||||||||||||||||||
Db  1267 ATTGTTCTGGCATATACTATGCCACCAAATGCTCAACATGTCACAGTTCTCAGAGTTGTC 1326

Qy   421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1327 ATTAGAGAAGATTTCTCCCGCACACTAGCGGAGCGACTGGTAATAGACATTGAAAAAGTC 1386

Qy   441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
                  :::|||   ||||||:::||||||      |||:::::::::
Db  1387 TTCCACGGAGTAGACACACTTCCGGCGAGGGTCAACGCTAAGCTAGCCGTG--------- 1437

Qy   461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
            :::|||:::|||         :::            ||||||:::|||   :::
Db  1438 ---GCCGAGGCGAATGGCAGCGGCGTGCAT---------AAGAAAACAGATAGAGAAGTG 1485

Qy   481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAsp---ArgLysLysThrSer 499
         |||   :::|||   |||   |||   |||||||||||||||   :::|||||||||:::
Db  1486 CAGCTAGAGATTACTACTGCATGGTTGAAATTTGTTGCTGATAAGAAGAAGAAGACTAAT 1545

Qy   500 GlyIleCys 502
         |||:::|||
Db  1546 GGAGTTTGT 1554

RESULT 15
AF353615
LOCUS       AF353615                1776 bp    mRNA    linear   PLN 02-MAY-2003
DEFINITION  Nicotiana tabacum glutamate decarboxylase isozyme 3 mRNA, complete
            cds.
ACCESSION   AF353615
VERSION     AF353615.1  GI:13752461
KEYWORDS    .
SOURCE      Nicotiana tabacum (common tobacco)
  ORGANISM  Nicotiana tabacum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; lamiids; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1776)
  AUTHORS   Yevtushenko,D., McLean,M.D., Peiris,S.E., Van Cauwenberghe,O.R. and
            Shelp,B.J.
  TITLE     Two isoforms of tobacco glutamate decarboxylase are regulated by
            calcium/calmodulin and differ in organ distribution
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1776)
  AUTHORS   McLean,M.D., Yevtushenko,D. and Shelp,B.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-2001) Department of Plant Agriculture, University
            of Guelph, Division of Biotechnology, Bovey Building, Guelph,
            Ontario N1G 2W1, Canada
FEATURES             Location/Qualifiers
     source          1..1776
                     /organism="Nicotiana tabacum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4097"
     CDS             89..1564
                     /EC_number="4.1.1.15"
                     /function="converts glutamate to gamma-aminobutyric acid"
                     /note="GAD; GAD4; NtGAD4"
                     /codon_start=1
                     /product="glutamate decarboxylase isozyme 3"
                     /protein_id="AAK38667.1"
                     /db_xref="GI:13752462"
                     /translation="MVLSKTSSESDVSVHSTFASRYVRTSLPRFEMAENSIPKEAAFQ
                     IINDELMLDGNPRLNLASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVN
                     MIARLFNAPLEEKETAVGVGTVGSSEAIMLAGLAFKRNWQNKRKAEGKPYNKPNIVTG
                     ANVQVCWEKFANYFEVELKEVKLREGYYVMDPVQAVEMVDENTICVAAILGSTLNGEF
                     EDVKLLNDLLIEKNKQTGWNTPIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGH
```

```
                     KYGLVYAGIGWVIWRTKQDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIR
                     LGYEGYRNVMENCRENAIVLREGLEKTGRFNIVSKDEGVPLVAFSLKDNSRHNEFEVS
                     ETLRRFGWIVPAYTMPADAQHVTVLRVVIREDFSRTLAERLVLDIVKVLHELDTLPAR
                     LSAKLEEVKLVKNGKKFELEVQREVTNYWKKFVLARKAPVC"
SE COUNT       520 a    343 c    418 g    491 t       4 others
IGIN

ignment Scores:
ed. No.:                 8.32e-176      Length:          1776
ore:                     2186.50        Matches:         421
rcent Similarity:        90.32%         Conservative:    27
st Local Similarity:     84.88%         Mismatches:      39
ery Match:               83.61%         Indels:          9
                         8              Gaps:            1

-10-006-852-2 (1-502) x AF353615 (1-1776)

          1 MetValLeuSerHisAlaValSerGluSerAspValSerValHisSerThrPheAlaSer 20
            ||||||||||||            ||||||||||||||||||||||||||||||||||||
         89 ATGGTTCTCTCGAAACCTCCTCTGAGTCGGACGTTTCGGTACACTCCACTTTTGCCTCT 148

         21 ArgTyrValArgThrSerLeuProArgPheLysMetProGluAsnSerIleProLysGlu 40
            ||||||||||||||||||||||||||||||:::|||   |||||||||||||||||||||
        149 CGCTATGTTCGAACTTCCCTTCCAAGGTTTGAGATGGCGGAGAATTCGATACCAAAAGAG 208

         41 AlaAlaTyrGlnIleIleAsnAspGluLeuMetLeuAspGlyAsnProArgLeuAsnLeu 60
            ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
        209 GCGGCATTTCAAATAATTAACGACGAGTTGATGCTTGACGGGAATCCAAGGCTGAACTTG 268

         61 AlaSerPheValThrThrTrpMetGluProGluCysAspLysLeuIleMetSerSerIle 80
            |||||||||||||||||||||||||||||||||||||||||||||:::|||   ||||||
        269 GCTTCATTTGTGACAACATGGATGGAGCCAGAGTGTGATAAGCTTATGATGGACTCCATT 328

         81 AsnLysAsnTyrValAspMetAspGluTyrProValThrThrGluLeuGlnAsnArgCys 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        329 AACAAGAACTATGTTGACATGGATGAATACCCTGTTACCACTGAGCTTCAGAATCGCTGC 388

        101 ValAsnMetIleAlaHisLeuPheAsnAlaProLeuGluGluAlaGluThrAlaValGly 120
            |||||||||||||||   ||||||||||||||||||||||||   |||||||||||||||
        389 GTGAACATGATAGCACGTTTATTCAACGCGCCACTAGAAGAGAAGGAGACAGCAGTTGGA 448

        121 ValGlyThrValGlySerSerGluAlaIleMetLeuAlaGlyLeuAlaPheLysArgLys 140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||   
        449 GTGGGTACAGTTGGTTCATCGGAGGCCATAATGCTAGCGGGGCTTGCATTCAAGAGAAAT 508

        141 TrpGlnAsnLysArgLysAlaGluGlyLysProValAspLysProAsnIleValThrGly 160
            |||||||||||||||||||||||||||||||||   :::|||||||||||||||||||||
        509 TGGCAAAACAAACGCAAAGCTGAGGGCAAACCTTACAATAAGCCCAACATTGTCACTGGC 568

        161 AlaAsnValGlnValCysTrpGluLysPheAlaArgTyrPheGluValGluLeuLysGlu 180
            |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
        569 GCCAATGTTCAGGTGTGCTGGGAGAAATTTGCCAACTATTTTGAAGTGGAATTGAAAGAA 628

        181 ValLysLeuSerGluGlyTyrTyrValMetAspProGlnGlnAlaValAspMetValAsp 200
            |||||||||   ||||||||||||||||||||||||   |||||||||:::|||||||||
        629 GTAAAGCTAAGGGAAGGGTACTATGTGATGGACCCAGTCCAGGCTGTGGAGATGGTTGAT 688

        201 GluAsnThrIleCysValAlaAspIleLeuGlySerThrLeuAsnGlyGluPheGluAsp 220
            |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
        689 GAGAACACCATTTGTGTTGCTGCAATCTTGGGTTCAACCCTTAATGGAGAATTTGAAGAT 748

        221 ValLysLeuLeuAsnAspLeuLeuValGluLysAsnLysGluThrGlyTrpAspThrPro 240
            ||||||||||||||||||||||||:::||||||||||||:::|||||||||:::||||||
        749 GTCAAGCTCTTGAATGATCTTTTGATTGAAAAGAACAAGCAAACTGGATGGAACACACCA 808

        241 IleHisValAspAlaAlaSerGlyGlyPheIleAlaProPheLeuTyrProGluLeuGlu 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        809 ATTCATGTGGATGCAGCAAGTGGTGGATTCATTGCACCATTCCTGTACCCAGAGCTGGAG 868

        261 TrpAspPheArgLeuProLeuValLysSerIleAsnValSerGlyHisLysTyrGlyLeu 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        869 TGGGACTTTAGGCTTCCCTTAGTGAAGAGCATAAATGTGAGTGGGCACAAATATGGGCTT 928

Qy          281 ValTyrAlaGlyIleGlyTrpValIleTrpArgAsnLysGluAspLeuProGluGluLeu 300
                |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db          929 GTCTATGCTGGTATTGGTTGGGTTATTTGGAGGACCAAACAAGATTTGCCTGAAGAACTC 988

Qy          301 IlePheHisIleAsnTyrLeuGlyAlaAspGlnProThrPheThrLeuAsnPheSerLys 320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          989 ATTTTCCACATCAACTATCTTGGAGCTGATCAGCCCACCTTTACTCTCAATTTCTCCAAA 1048

Qy          321 GlySerSerGlnValIleAlaGlnTyrTyrGlnLeuIleArgLeuGlyHisGluGlyTyr 340
                ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db         1049 GGTTCAAGTCAAGTCATTGCTCAATATTATCAGCTAATCCGTTTGGGCTATGAGGGGTAC 1108

Qy          341 ArgAsnValMetGluAsnCysArgGluAsnMetIleValLeuArgGluGlyLeuGluLys 360
                ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db         1109 CGAAATGTAATGGAGAACTGTCGCGAAAATGCCATTGTGCTAAGAGAAGGACTCGAAAAA 1168

Qy          361 ThrGluArgPheAsnIleValSerLysAspGluGlyValProLeuValAlaPheSerLeu 380
                |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1169 ACAGGACGTTTTAACATAGTCTCCAAAGATGAAGGTGTCCCTTTGGTGGCCTTTTCCCTC 1228

Qy          381 LysAspSerSerCysHisThrGluPheGluIleSerAspMetLeuArgArgTyrGlyTrp 400
                ||||||:::|||   |||   |||||||||:::|||:::   |||||||||:::||||||
Db         1229 AAGGACAATAGCCGTCACAATGAGTTCGAGGTGTCCGAGACGCTCCGTAGGTTCGGGTGG 1288

Qy          401 IleValProAlaTyrThrMetProProAsnAlaGlnHisIleThrValLeuArgValVal 420
                ||||||||||||||||||||||||   :::|||||||||:::||||||||||||||||||
Db         1289 ATCGTCCCGGCCTACACGATGCCCGCTGACGCCCAACACGTCACGGTGCTTCGTGTGGTG 1348

Qy          421 IleArgGluAspPheSerArgThrLeuAlaGluArgLeuValIleAspIleGluLysVal 440
                ||||||||||||||||||||||||||||||||||||||||||:::||||||   ||||||
Db         1349 ATCCGGGAGGACTTCTCGCGAACCCTAGCAGAGCGTCTCGTCCTCGACATTGTCAAGGTC 1408

Qy          441 MetArgGluLeuAspGluLeuProSerArgValIleHisLysIleSerLeuGlyGlnGlu 460
                :::   |||||||||   ||||||:::|||:::      |||:::
Db         1409 CTCCACGAGCTGGACACACTTCCAGCTAGGCTGAGCGCCAAATTA--------------- 1453

Qy          461 LysSerGluSerAsnSerAspAsnLeuMetValThrValLysLysSerAspIleAspLys 480
                                        |||:::         ||||||   :::::::::
Db         1454 ------------GAGGAGGTGAAGCTGGTCAAGAATGGAAAGAAATTTGAACTTGAAGTT 1501

Qy          481 GlnArgAspIleIleThrGlyTrpLysLysPheValAlaAspArgLys 496
                ||||||::::::         |||||||||||||||      ||||||
Db         1502 CAAAGGGAAGTTACCAATTATTGGAAGAAGTTTGTTTTAGCTAGGAAA 1549

Search completed: October 22, 2003, 16:47:12
Job time : 4597 secs
```

```
                  GenCore version 5.1.6
         Copyright (c) 1993 - 2003  Compugen Ltd.

M protein - protein search, using sw model

n on:         October 22, 2003, 15:07:49 ; Search time 443 Seconds
                                          (without alignments)
                                          189.764 Million cell updates/sec

tle:          US-10-006-852-2
rfect score:  2615
quence:       1 MVLSHAVSESDVSVHSTFAS..........DIITGWKKFVADRKKTSGIC 502

oring table:  BLOSUM62
              Gapop 10.0 , Gapext 0.5

arched:       629382 seqs, 167460630 residues

tal number of hits satisfying chosen parameters:      629382

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
               Maximum Match 100%
               Listing first 45 summaries

tabase :       Published_Applications_AA:*
              1:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
              2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
              3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
              4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
              5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
              6:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
              7:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
              8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
              9:  /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
              10:  /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
              11:  /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
              12:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
              13:  /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
              14:  /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
              15:  /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
              16:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
              17:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
              18:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                                  SUMMARIES
                      %
esult             Query
 No.    Score  Match Length DB  ID                          Description
------------------------------------------------------------------------------
   1     2615  100.0    502  15  US-10-006-852-2              Sequence 2, Appli
   2     2250   86.0    500  15  US-10-006-852-16             Sequence 16, Appl
   3     2240   85.7    496  15  US-10-006-852-12             Sequence 12, Appl
   4     2226   85.1    496  15  US-10-005-602-2              Sequence 2, Appli
   5     2199   84.1    496  15  US-10-006-852-14             Sequence 14, Appl
   6   2135.5   81.7    493  15  US-10-006-852-8              Sequence 8, Appli
   7     2130   81.5    494  15  US-10-006-852-4              Sequence 4, Appli
   8     2080   79.5    500  15  US-10-006-852-6              Sequence 6, Appli
   9     2073   79.3    498  12  US-10-167-547C-2             Sequence 2, Appli
  10     2014   77.0    502  15  US-10-006-852-18             Sequence 18, Appl
  11     1969   75.3    494  15  US-10-006-852-10             Sequence 10, Appl
  12     1864   71.3    509  12  US-10-167-547C-4             Sequence 4, Appli
  13     1864   71.3    529  12  US-10-167-547C-6             Sequence 6, Appli
  14     1082   41.4    470  15  US-10-156-761-12189          Sequence 12189, A
  15     1051   40.2    454  15  US-10-156-761-11138          Sequence 11138, A
```

```
  16   393.5   15.0    191  11  US-09-989-442-145           Sequence 145, App
  17     325   12.4    589  12  US-10-197-073-8             Sequence 8, Appli
  18     325   12.4    589  12  US-10-053-510-2             Sequence 2, Appli
  19     325   12.4    589  15  US-10-286-175-8             Sequence 8, Appli
  20   299.5   11.5    545  12  US-10-053-510-16            Sequence 16, Appl
  21     288   11.0    552  12  US-10-053-510-11            Sequence 11, Appl
  22   264.5   10.1    542  12  US-10-197-073-6             Sequence 6, Appli
  23   264.5   10.1    542  12  US-10-053-510-4             Sequence 4, Appli
  24   264.5   10.1    542  15  US-10-286-175-6             Sequence 6, Appli
  25   246.5    9.4    568  12  US-10-197-073-2             Sequence 2, Appli
  26   246.5    9.4    568  12  US-10-053-510-6             Sequence 6, Appli
  27   246.5    9.4    568  15  US-10-286-175-2             Sequence 2, Appli
  28     227    8.7    568  12  US-10-197-073-4             Sequence 4, Appli
  29     227    8.7    568  12  US-10-053-510-8             Sequence 8, Appli
  30     227    8.7    568  15  US-10-286-175-4             Sequence 4, Appli
  31     224    8.6    568  10  US-09-740-369-2             Sequence 2, Appli
  32     224    8.6    568  12  US-10-053-510-18            Sequence 18, Appl
  33   160.5    6.1    488  12  US-10-197-073-10            Sequence 10, Appl
  34   160.5    6.1    488  12  US-10-053-510-10            Sequence 10, Appl
  35   160.5    6.1    488  15  US-10-286-175-10            Sequence 10, Appl
  36   118.5    4.5    384  15  US-10-156-761-10470         Sequence 10470, A
  37     116    4.4    509   9  US-09-846-590B-20          Sequence 20, Appl
  38   109.5    4.2    946  12  US-10-032-585-7900          Sequence 7900, Ap
  39   106.5    4.1    372  11  US-09-765-061B-74           Sequence 74, Appl
  40   106.5    4.1    392  11  US-09-765-061B-77           Sequence 77, Appl
  41   105.5    4.0   1320  12  US-10-161-051-167           Sequence 167, App
  42   105.5    4.0   1425  12  US-10-168-424-2             Sequence 2, Appli
  43   103.5    4.0    525   9  US-09-815-242-11741        Sequence 11741, A
  44   101.5    3.9    820   9  US-09-815-242-5882         Sequence 5882, Ap
  45   101.5    3.9    825   9  US-09-815-242-13079        Sequence 13079, A

                                  ALIGNMENTS


RESULT 1
US-10-006-852-2
; Sequence 2, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 502
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-006-852-2

  Query Match             100.0%;  Score 2615;  DB 15;  Length 502;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-245;
  Matches  502;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240

            241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300

            301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360

            361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420

            421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
            421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480

            481 QRDIITGWKKFVADRKKTSGIC 502
                ||||||||||||||||||||||
            481 QRDIITGWKKFVADRKKTSGIC 502

SULT 2
-10-006-852-16
Sequence 16, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
 APPLICANT: Kinnersely, Alan M.
 APPLICANT:  Turano, Frank J.
 TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 FILE REFERENCE: 7224-65
 CURRENT APPLICATION NUMBER: US/10/006,852
 CURRENT FILING DATE:  2002-07-01
 PRIOR APPLICATION NUMBER: US 60/246,367
 PRIOR FILING DATE: 2000-11-07
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
  LENGTH: 500
  TYPE: PRT
  ORGANISM: Petunia x hybrida
-10-006-852-16

Query Match            86.0%;  Score 2250;  DB 15;  Length 500;
Best Local Similarity   84.9%;  Pred. No. 2.3e-209;
Matches  427;  Conservative   39;  Mismatches   33;  Indels    4;  Gaps    2;

y             1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
                ||||  ||:||||:|||||||||||||||||||:||||||||||||||||||||||||||
b             1 MVLSKTVSQSDVSIHSTFASRYVRTSLPRFKMPDNSIPKEAAYQIINDELMLDGNPRLNL 60

y            61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
                |||||||||||||||:| |||||||||||||||||||||||||||||||||||: |||||
b            61 ASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEDGETAVG 120

y           121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
                |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
b           121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTGANVQVCWEKFARYFEVELKE 180

y           181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
                ||||||||||||::||:|||||||||| |||||||||||||| |||||||||||||||||
b           181 VKLSEGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVEKNKETGWDTP 240

y           241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
                ||||||||||||||:|||||||||||||||||||||||||||||||||:||||:|||:||
b           241 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRNKDDLPDEL 300

y           301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
                ||||||||||||||||||||||||||||||||||||:|||:||||||:||  ||||||||
b           301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGYEGYKNVMENCQENASVLREGLEK 360
```

```
Qy          361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
                | ||||:||: |||||||||||:  | |||||: |||:||||||||||||||||||||||
Db          361 TGRFNIISKEIGVPLVAFSLKDNRQHNEFEISETLRRFGWIVPAYTMPPNAQHITVLRVV 420

Qy          421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
                |||||||||||||| |||||: ||| ||:||  |::: :|:: :|   :    ||:| : 
Db          421 IREDFSRTLAERLVRDIEKVLHELDTLPARVNAKLAVAEEQAAANGSEVH---KKTDSEV 477

Qy          481 QRDIITGWKKFVAD-RKKTSGIC 502
                | ::|| ||||| : :|||: :|
Db          478 QLEMITAWKKFVEEKKKKTNRVC 500

RESULT 3
US-10-006-852-12
; Sequence 12, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
;   LENGTH: 496
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-10-006-852-12

  Query Match            85.7%;  Score 2240;  DB 15;  Length 496;
  Best Local Similarity   85.1%;  Pred. No. 2.1e-208;
  Matches  428;  Conservative   35;  Mismatches   32;  Indels    8;  Gaps    3;

Qy            1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
                ||||   ||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db            1 MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy           61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
                ||||||||||||:||:| |||||||||||||||||||||||||||||||||| : |||||
Db           61 ASFVTTWMEPECNKLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

Qy          121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
                |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
Db          121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy          181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
                ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
Db          181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

Qy          241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
                ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||:||
Db          241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

Qy          301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
                |||||||||||||||||||||||||||||||||||| |||:||||||:||  ||||||||
Db          301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGLEK 360

Qy          361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
                : ||||:||: |||||||||||:| | |||||: |||:|||:|||||||||||:||||||
Db          361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIIPAYTMPPNAQHVTVLRVV 420

Qy          421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
                ||||||||||||||||||||: ||| ||:||  |:::    :|:|   :    ||:| : 
Db          421 IREDFSRTLAERLVIDIEKVLHELDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

Qy          481 QRDIITGWKKFVAD-RKKTSGIC 502
                | :| | ||||||| :|||:|:|
Db          474 QLEITTAWKKFVADKKKKTNGVC 496
```

```
ESULT 4
S-10-005-602-2
 Sequence 2, Application US/10005602
 Publication No. US20030110530A1
 GENERAL INFORMATION:
  APPLICANT: Barry Shelp
  APPLICANT:  Alan Bown
  TITLE OF INVENTION: TRANSGENIC PLANTS HAVING REDUCED
  TITLE OF INVENTION:  SUSCEPTIBILITY TO INVERTEBRATE PESTS
  FILE REFERENCE: P84US3
  CURRENT APPLICATION NUMBER: US/10/005,602
  CURRENT FILING DATE:  2002-12-31
  NUMBER OF SEQ ID NOS: 5
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 2
   LENGTH: 496
   TYPE: PRT
   ORGANISM: tobacco plant
-10-005-602-2

Query Match             85.1%;  Score 2226;  DB 15;  Length 496;
Best Local Similarity   84.7%;  Pred. No. 4.9e-207;
Matches  426;  Conservative   35;  Mismatches   34;  Indels    8;  Gaps    3;

           1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
             ||||   ||||||:||||||||||||||||||||||||||||||||||||||||||||||
           1 MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

          61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
             ||||||||||||: |:| |||||||||||||||||||||||||||||||||| : |||||
          61 ASFVTTWMEPECNTLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
             |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
         121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPFDKPNIVTGANVQVCWEKFARYFEVELKE 180

         181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
             ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
         181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
             ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||:||
         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
             |||||||||||||||||||||||||||||||||||| |||:||||||:||  |||||:||
         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGIEK 360

         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
             : ||||:||: |||||||||||:| | |||||: |||:||||||||||||| |:||||||
         361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIVPAYTMPPNAGHVTVLRVV 420

         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
             ||||||||||||||||||||: ||| ||:||  |:::    :|:|   :    ||:| :
         421 IREDFSRTLAERLVIDIEKVLHELDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

         481 QRDIITGWKKFVAD-RKKTSGIC 502
             | :| | ||||||| :|||:|:|
         474 QLEITTAWKKFVADKKKKTNGVC 496

ESULT 5
S-10-006-852-14
 Sequence 14, Application US/10006852
 Publication No. US20030046732A1
 GENERAL INFORMATION:
  APPLICANT: Kinnersely, Alan M.
  APPLICANT:  Turano, Frank J.
  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
  FILE REFERENCE: 7224-65
  CURRENT APPLICATION NUMBER: US/10/006,852
```

```
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 496
;   TYPE: PRT
;   ORGANISM: Nicotiana tabacum
US-10-006-852-14

  Query Match             84.1%;  Score 2199;  DB 15;  Length 496;
  Best Local Similarity   84.1%;  Pred. No. 2e-204;
  Matches  423;  Conservative   34;  Mismatches   38;  Indels    8;  Gaps    3;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVLSKTASESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||: |:| |||||||||||||||||||||||||||||||||| : |||||
Db         61 ASFVTTWMEPECNTLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPFDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
Db        181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||| ||||||||||||||||||| |||||||||:||
Db        241 IHVDAASGGFIAPFLYPELEWDFRLPLEKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||:||||||:||  |||||:||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGIEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              : ||||:||: |||||||||||:| | |||||: |||:|||| ||||||||||:||||||
Db        361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIVLAYTMPPNAQHVTVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||||||||   :| ||:||  |:::    :|:|   :    ||:| :
Db        421 IREDFSRTLAERLVIDIEKVFHGVDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

Qy        481 QRDIITGWKKFVAD-RKKTSGIC 502
              | :| | | ||||| :|||:|:|
Db        474 QLEITTAWLKFVADKKKKTNGVC 496


RESULT 6
US-10-006-852-8
; Sequence 8, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 493
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-006-852-8
```

```
Query Match             81.7%;  Score 2135.5;  DB 15;  Length 493;
Best Local Similarity   82.4%;  Pred. No. 2.9e-198;
Matches  411;  Conservative   32;  Mismatches   45;  Indels   11;  Gaps    2;

          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
            ||||  |||||||:|||||||||| |||||:|||||||||||||||||||||||||||||
          1 MVLSKTVSESDVSIHSTFASRYVRNSLPRFEMPENSIPKEAAYQIINDELMLDGNPRLNL 60

         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
            |||||||||||||||:| ||||||||||||||||||||||||||| |||||| : | |||
         61 ASFVTTWMEPECDKLMMESINKNYVDMDEYPVTTELQNRCVNMIARLFNAPLGDGEAAVG 120

        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
            |||||||||||||||||||:|||||||:| | ||||||||||||||||||||||||||||
        121 VGTVGSSEAIMLAGLAFKRQWQNKRKAQGLPYDKPNIVTGANVQVCWEKFARYFEVELKE 180

        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
            | | | |||||| :||:|||||||||| |||||| ||||||||||||||||||:||||||
        181 VNLREDYYVMDPVKAVEMVDENTICVAAILGSTLTGEFEDVKLLNDLLVEKNKQTGWDTP 240

        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
            ||||||||||||||||||||||||||||||||||||||||||||||||:|| | |||:||
        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
            |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
            | || ||||: ||||||||||||| | |||::  |||:|||||||||| :|||:||||||
        361 TGRFKIVSKENGVPLVAFSLKDSSRHNEFEVAHTLRRFGWIVPAYTMPADAQHVTVLRVV 420

        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
            |||||||||||||| | |||: ||| ||:||  |:: |:             |||:  :
        421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMANGKVNG----------VKKTPEET 470

        481 QRDIITGWKKFVADRKKTS 499
            ||::   ||| : : |||:
        471 QREVTAYWKKLL-ETKKTN 488
```

SULT 7
-10-006-852-4
Sequence 4, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
 APPLICANT: Kinnersely, Alan M.
 APPLICANT: Turano, Frank J.
 TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 FILE REFERENCE: 7224-65
 CURRENT APPLICATION NUMBER: US/10/006,852
 CURRENT FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: US 60/246,367
 PRIOR FILING DATE: 2000-11-07
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
  LENGTH: 494
  TYPE: PRT
  ORGANISM: Arabidopsis thaliana
S-10-006-852-4

```
 Query Match             81.5%;  Score 2130;  DB 15;  Length 494;
 Best Local Similarity   80.2%;  Pred. No. 1e-197;
 Matches  404;  Conservative   46;  Mismatches   42;  Indels   12;  Gaps    3;

          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
            |||:   : :| || : | ||||||:||:::: ||||||:|||||| |||||||||||||
          1 MVLTKTAT-NDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQIIKDELMLDGNPRLNL 59

         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
            ||||||||||||||||| ||||||||||||||||||||||||:|| ||||||||:|||||
         60 ASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNIIARLFNAPLEESETAVG 119
```

```
Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        120 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGANVQVCWEKFARYFEVELKE 179

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              | |||||||||| :| :|||||||||| |||||||||||||| ||||||:||:||||:||
Db        180 VNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVKKNEETGWNTP 239

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||:|||||||||||||||||||||||||||||||||:||  |||||||
Db        240 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRAAEDLPEEL 299

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              ||||||||||||||||||||||||:||||||||||| |||:|||||| |||:||:||:||
Db        300 IFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRLGFEGYKNVMENCIENMVVLKEGIEK 359

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              ||||||||||:|||:||||||| | | |||||:||||:|||||||||| :||||||||||
Db        360 TERFNIVSKDQGVPVVAFSLKDHSFHNEFEISEMLRRFGWIVPAYTMPADAQHITVLRVV 419

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| || ||: ||| |||::  |: :         : :   ||:  ::|
Db        420 IREDFSRTLAERLVADISKVLHELDTLPSKISKKMGI---------EGIAENVKEKKMEK 470

Qy        481 Q--RDIITGWKKFVADRKKTSGIC 502
              :   ::| ||:||| :||| :|:|
Db        471 EILMEVIVGWRKFVKERKKMNGVC 494
```

RESULT 8
US-10-006-852-6
; Sequence 6, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
; APPLICANT: Kinnersely, Alan M.
; APPLICANT: Turano, Frank J.
; TITLE OF INVENTION: Methods for Regulating Plant GABA Production
; FILE REFERENCE: 7224-65
; CURRENT APPLICATION NUMBER: US/10/006,852
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 60/246,367
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 500
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-006-852-6

```
  Query Match             79.5%;  Score 2080;  DB 15;  Length 500;
  Best Local Similarity   78.8%;  Pred. No. 7.3e-193;
  Matches  395;  Conservative   44;  Mismatches   52;  Indels   10;  Gaps    1;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   |:|| |:|||||||||| |: ||::|:||||||||||||||||  ||||||||
Db          1 MVLSKTASKSDDSIHSTFASRYVRNSISRFEIPKNSIPKEAAYQIINDELKFDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              |||||||||||||||:| ||||| |:||:|||||:|||||||||| |||||| : | |:|
Db         61 ASFVTTWMEPECDKLMMESINKNNVEMDQYPVTTDLQNRCVNMIARLFNAPLGDGEAAIG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||:|||||||||:||||||| | | |:||||||||:||| |||||||||||||
Db        121 VGTVGSSEAVMLAGLAFKRQWQNKRKALGLPYDRPNIVTGANIQVCLEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||| |||||||| :||:|||||||||  |||||| ||||||||||||||||||:||||||
Db        181 VKLREGYYVMDPDKAVEMVDENTICVVAILGSTLTGEFEDVKLLNDLLVEKNKKTGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              |||||||||||||||||:||||||||||||||||||||||||||||||:|| | |||:||
```

```
     241 IHVDAASGGFIAPFLYPDLEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

     301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
         |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
     301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

     361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
         | |||||||: ||||||||||||| | |||:::||||:|||||||||| :|||:||||||
     361 TGRFNIVSKENGVPLVAFSLKDSSRHNEFEVAEMLRRFGWIVPAYTMPADAQHVTVLRVV 420

     421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
         |||||||||||||| | |||: ||| ||:||  |:: |:             |||:  :
     421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMASGKVNG----------VKKTPEET 470

     481 QRDIITGWKKFVADRKKTSGI 501
         ||::   |||||  :   :|:
     471 QREVTAYWKKFVDTKTDKNGV 491


SULT 9
-10-167-547C-2
Sequence 2, Application US/10167547C
Publication No. US20030170653A1
GENERAL INFORMATION:
 APPLICANT: E.I. du Pont de Nemours and Company
 APPLICANT:  Damude, Howard G.
 TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
 TITLE OF INVENTION:  Butyrolactone and its Intermediates
 FILE REFERENCE: CL1804 US NA
 CURRENT APPLICATION NUMBER: US/10/167,547C
 CURRENT FILING DATE:  2003-03-17
 PRIOR APPLICATION NUMBER: 60/297198
 PRIOR FILING DATE: 2001-06-08
 NUMBER OF SEQ ID NOS: 67
 SOFTWARE: Microsoft Office 07
SEQ ID NO 2
  LENGTH: 498
  TYPE: PRT
  ORGANISM: Alstroemeria
-10-167-547C-2

Query Match            79.3%;  Score 2073;  DB 12;  Length 498;
Best Local Similarity  77.4%;  Pred. No. 3.5e-192;
Matches  391;  Conservative   53;  Mismatches   51;  Indels   10;  Gaps    2;

           1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
             |||| |||::   || |||||||| :  ||:||||||||: ||||:||||||||||||||
           1 MVLSSAVSDNTGPVHCTFASRYVRDAPARFRMPENSIPKDTAYQIVNDELMLDGNPRLNL 60

          61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
             |||||||||||||:|:::: ||||||||||||||||||||||:||||||||: : |||||
          61 ASFVTTWMEPECDRLMIAAANKNYVDMDEYPVTTELQNRCVNIIAHLFNAPIGDEETAVG 120

         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
             ||||||||||||||||||||||||:|||||| ||||||||||||||||||||||||||||
         121 VGTVGSSEAIMLAGLAFKRKWQNKKKAEGKPYDKPNIVTGANVQVCWEKFARYFEVELKE 180

         181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
             ||| ||||:|||::||:|||||||||| |||||| |||||||||| || |||||||||||
         181 VKLREGYYIMDPEKAVEMVDENTICVAAILGSTLTGEFEDVKLLNKLLEEKNKETGWDTP 240

         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
             ||||||||||||||||||||||||||||||||||||||||||||:|||:|||||||||||
         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGVGWVVWRNKEDLPEEL 300

         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
             ||||||||||||||||||||||||:|||||| |||| :||:|:||||  ||  :|||||::
         301 IFHINYLGADQPTFTLNFSKGSSQIIAQYYQFIRLGFQGYKNIMENCMENTRILREGLQE 360

         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
             | || ||||| |||||||:||||| :| |||:| :||:|||:|||||| :|:|| |||||
         361 TGRFEIVSKDIGVPLVAFALKDSSQYTVFEIADAMRRFGWIIPAYTMPKDAEHIAVLRVV 420
```

```
Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSR---VIHKISLGQEKSESNSDNLMVTVKKSD 477
              |||||||:|||||| |::||: ||| ||||   : |  ::        :||    :|||
Db        421 IREDFSRSLAERLVNDMKKVLAELDVLPSRITTIAHVTAV-------ENDNGEAVIKKSF 473

Qy        478 IDKQRDIITGWKKFVADRKKTSGIC 502
              :: :: :|| ||  | : |||: :|
Db        474 LEIEKKVITHWKDVVMNGKKTNKVC 498


RESULT 10
US-10-006-852-18
; Sequence 18, Application US/10006852
; Publication No. US20030046732A1
; GENERAL INFORMATION:
;  APPLICANT: Kinnersely, Alan M.
;  APPLICANT:  Turano, Frank J.
;  TITLE OF INVENTION: Methods for Regulating Plant GABA Production
;  FILE REFERENCE: 7224-65
;  CURRENT APPLICATION NUMBER: US/10/006,852
;  CURRENT FILING DATE:  2002-07-01
;  PRIOR APPLICATION NUMBER: US 60/246,367
;  PRIOR FILING DATE: 2000-11-07
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
;   LENGTH: 502
;   TYPE: PRT
;   ORGANISM: Lycopersicon esculentum
US-10-006-852-18

  Query Match            77.0%;  Score 2014;  DB 15;  Length 502;
  Best Local Similarity  76.0%;  Pred. No. 1.9e-186;
  Matches  387;  Conservative   53;  Mismatches   55;  Indels   14;  Gaps    5;

Qy          1 MVL-SHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLN 59
              ||| : :: :|: |:| |||||||:  ||:||||: |:||||||||:|||||||||||||
Db          1 MVLTTTSIRDSEESLHCTFASRYVQEPLPKFKMPKKSMPKEAAYQIVNDELMLDGNPRLN 60

Qy         60 LASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAV 119
              |||||:||||||||||||||||||||||||||||||||||||||:||||:||: :  ||||
Db         61 LASFVSTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMLAHLFHAPVGDDETAV 120

Qy        120 GVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELK 179
              |||||||||||||||||||||||:|||||||| |||||||||||||||||||||||||||
Db        121 GVGTVGSSEAIMLAGLAFKRKWQSKRKAEGKPFDKPNIVTGANVQVCWEKFARYFEVELK 180

Qy        180 EVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDT 239
              |||| |||||||| :||::||||||||| |||||| ||||||||||:|| :|||||||:|
Db        181 EVKLKEGYYVMDPAKAVEIVDENTICVAAILGSTLTGEFEDVKLLNELLTKKNKETGWET 240

Qy        240 PIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEE 299
              ||||||||||||||||:|:||||||||||||||||||||||||||:||||||:|||||:|
Db        241 PIHVDAASGGFIAPFLWPDLEWDFRLPLVKSINVSGHKYGLVYAGVGWVIWRSKEDLPDE 300

Qy        300 LIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLE 359
              |:|||||||:||||||||||||| |:||||||||||| |||:|||:||  |  || ||:
Db        301 LVFHINYLGSDQPTFTLNFSKGSYQIIAQYYQLIRLGFEGYKNVMKNCLSNAKVLTEGIT 360

Qy        360 KTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRV 419
              |  ||:||||| |||:|||||:||| :| ||:|:| |||:|||||||||||:|:|| ||||
Db        361 KMGRFDIVSKDVGVPVVAFSLRDSSKYTVFEVSEHLRRFGWIVPAYTMPPDAEHIAVLRV 420

Qy        420 VIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKI--SLGQEKSESNSDNL----MVTV 473
              ||||||| :|||||| ||||:: |||  | |:  |      :|  :   | |    | ||
Db        421 VIREDFSHSLAERLVSDIEKILSELDTQPPRLPTKAVRVTAEEVRDDKGDGLHHFHMDTV 480

Qy        474 KKSDIDKQRDIITGWKKFVADRKKTSGIC 502
                   : |:|||  |:|     |||||:|
Db        481 -----ETQKDIIKHWRKIAG--KKTSGVC 502


RESULT 11
US-10-006-852-10
```

```
Sequence 10, Application US/10006852
Publication No. US20030046732A1
GENERAL INFORMATION:
 APPLICANT: Kinnersely, Alan M.
 APPLICANT:  Turano, Frank J.
 TITLE OF INVENTION: Methods for Regulating Plant GABA Production
 FILE REFERENCE: 7224-65
 CURRENT APPLICATION NUMBER: US/10/006,852
 CURRENT FILING DATE:  2002-07-01
 PRIOR APPLICATION NUMBER: US 60/246,367
 PRIOR FILING DATE: 2000-11-07
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
  LENGTH: 494
  TYPE: PRT
  ORGANISM: Arabidopsis thaliana
-10-006-852-10

Query Match             75.3%;  Score 1969;  DB 15;  Length 494;
Best Local Similarity   75.2%;  Pred. No. 4.2e-182;
Matches  373;  Conservative   56;  Mismatches   61;  Indels    6;  Gaps    3;

      1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
        |||:   |:||  :||||||||||  :||||||:: :||:||||:|||||||||||||||
      1 MVLA-TNSDSDEHLHSTFASRYVRAVVPRFKMPDHCMPKDAAYQVINDELMLDGNPRLNL 59

     61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
        ||||||||||||||||| |:|||||||||||||||||||||||||:||:||: | | |:|
     60 ASFVTTWMEPECDKLIMDSVNKNYVDMDEYPVTTELQNRCVNMIANLFHAPVGEDEAAIG 119

    121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
         |||||||||||||||||||||::|||:| |:||||||||||||||||||||||||||||
    120 CGTVGSSEAIMLAGLAFKRKWQHRRKAQGLPIDKPNIVTGANVQVCWEKFARYFEVELKE 179

    181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
        ||||| |||||| :||:|||||||||| |||||| ||||||| ||||| ||| ||||:||
    180 VKLSEDYYVMDPAKAVEMVDENTICVAAILGSTLTGEFEDVKQLNDLLAEKNAETGWETP 239

    241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
        |||||||||||||||||:|||||||| |||||||||||||||||:|||:|| |:||||||
    240 IHVDAASGGFIAPFLYPDLEWDFRLPWVKSINVSGHKYGLVYAGVGWVVWRTKDDLPEEL 299

    301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
        :|||||||||||||||||||||||:|||||| |||| |||:|:|||| :|   ||||:|
    300 VFHINYLGADQPTFTLNFSKGSSQIIAQYYQFIRLGFEGYKNIMENCMDNARRLREGIEM 359

    361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
        | :||||||| ||||||||||||| || |||:: ||::|||:|||||| :|||| |||||
    360 TGKFNIVSKDIGVPLVAFSLKDSSKHTVFEIAESLRKFGWIIPAYTMPADAQHIAVLRVV 419

    421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
        ||||||| ||:||:  | :|::|:: ||||: |  :     :  : |:  | || : :
    420 IREDFSRGLADRLITHIIQVLKEIEGLPSRIAHLAA----AAAVSGDDEEVKVKTAKMSL 475

    481 QRDIITGWKKFVADRK 496
        : ||   ||: |  ::
    476 E-DITKYWKRLVEHKR 490

ESULT 12
S-10-167-547C-4
 Sequence 4, Application US/10167547C
 Publication No. US20030170653A1
 GENERAL INFORMATION:
  APPLICANT: E.I. du Pont de Nemours and Company
  APPLICANT:  Damude, Howard G.
  TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
  TITLE OF INVENTION:  Butyrolactone and its Intermediates
  FILE REFERENCE: CL1804 US NA
  CURRENT APPLICATION NUMBER: US/10/167,547C
  CURRENT FILING DATE:  2003-03-17
  PRIOR APPLICATION NUMBER: 60/297198
```

```
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: Microsoft Office 07
; SEQ ID NO 4
;   LENGTH: 509
;   TYPE: PRT
;   ORGANISM: Alstroemeria
US-10-167-547C-4

  Query Match             71.3%;  Score 1864;  DB 12;  Length 509;
  Best Local Similarity   73.3%;  Pred. No. 6.9e-172;
  Matches  351;  Conservative   54;  Mismatches   70;  Indels    4;  Gaps    2;

Qy        1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
            | ||  ||:|:  |  |:||||||   | |:||| |||||||: :|||||||||||||||
Db        1 MALSSVVSDSNNQVQCTYASRYVRDEAPGFRMPEKSIPKEAAFTMINDELMLDGNPRLNL 60

Qy       61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
            |||||||||||||:|:||:|||||  ||:|||| ::|||||||||:|||||: | || ||
Db       61 ASFVTTWMEPECDRLMMSTINKNYALMDDYPVTIDIQNRCVNMIANLFNAPIGEGETTVG 120

Qy      121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||:||||||||| ||||||||||| ||||:|||:|||||| |||:|||||:|:
Db      121 CATVGSSEAMMLAGLAFKRNWQNKRKAEGKPYDKPNMVTGSNVQVCWVKFAKYFEVEMKK 180

Qy      181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
            | | ||||||||::||:|||||||||| |||||| ||||||||||||||||||:||||||
Db      181 VNLREGYYVMDPEKAVEMVDENTICVAAILGSTLTGEFEDVKLLNDLLVEKNKKTGWDTP 240

Qy      241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
            |||||| |||||||:||||||||||||||||||||||||||| |:|||:|||| ||||||
Db      241 IHVDAAIGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYPGVGWVVWRNKNDLPEEL 300

Qy      301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
            |||||||| |||||||||||||:|:| ||||||||| |||:|:|||| ||  :||| ||:
Db      301 IFHINYLGIDQPTFTLNFSKGSNQIIGQYYQLIRLGFEGYKNIMENCTENARILREHLEE 360

Qy      361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
               | |:||| | |||  :||||| |: |:|:| :||:|| :|||||| : :|| |||||
Db      361 MGVFEIISKDIGAPLVTIALKDSSKHSVFKIADTIRRFGWTIPAYTMPKDVEHIAVLRVV 420

Qy      421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDID 479
            |||||||:|||||  |::||: |||  |||     ::   |:  | :   | :||| ::
Db      421 IREDFSRSLAERLANDMKKVLVELDIHPSRT---TTIAHVKAVENGNGNYV-IKKSIVE 475

RESULT 13
US-10-167-547C-6
; Sequence 6, Application US/10167547C
; Publication No. US20030170653A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. du Pont de Nemours and Company
;  APPLICANT:  Damude, Howard G.
;  TITLE OF INVENTION: A Biological Method for the Production of Alpha-Methylene-Gamma
;  TITLE OF INVENTION:  Butyrolactone and its Intermediates
;  FILE REFERENCE: CL1804 US NA
;  CURRENT APPLICATION NUMBER: US/10/167,547C
;  CURRENT FILING DATE:  2003-03-17
;  PRIOR APPLICATION NUMBER: 60/297198
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: Microsoft Office 07
; SEQ ID NO 6
;   LENGTH: 529
;   TYPE: PRT
;   ORGANISM: Alstroemeria
US-10-167-547C-6

  Query Match             71.3%;  Score 1864;  DB 12;  Length 529;
  Best Local Similarity   73.3%;  Pred. No. 7.3e-172;
  Matches  351;  Conservative   54;  Mismatches   70;  Indels    4;  Gaps    2;

Qy        1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
```

```
     | ||  ||:|:  |  |:||||||   | |:||| |||||||: :|||||||||||||||
   1 MALSSVVSDSNNQVQCTYASRYVRDEAPGFRMPEKSIPKEAAFTMINDELMLDGNPRLNL 60

  61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
     |||||||||||||:|:||:|||||  ||:|||| ::|||||||||:|||||: | || ||
  61 ASFVTTWMEPECDRLMMSTINKNYALMDDYPVTIDIQNRCVNMIANLFNAPIGEGETTVG 120

 121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
       |||||||:||||||||| ||||||||||| ||||:|||:|||||| |||:|||||:|:
 121 CATVGSSEAMMLAGLAFKRNWQNKRKAEGKPYDKPNMVTGSNVQVCWVKFAKYFEVEMKK 180

 181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
     | | ||||||||::||:|||||||||| |||||| ||||||||||||||||||:||||||
 181 VNLREGYYVMDPEKAVEMVDENTICVAAILGSTLTGEFEDVKLLNDLLVEKNKKTGWDTP 240

 241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
     |||||| |||||||:||||||||||||||||||||||||||| |:|||:|||| ||||||
 241 IHVDAAIGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYPGVGWVVWRNKNDLPEEL 300

 301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
     |||||||| |||||||||||||:|:| ||||||||| |||:|:|||| ||  :||| ||:
 301 IFHINYLGIDQPTFTLNFSKGSNQIIGQYYQLIRLGFEGYKNIMENCTENARILREHLEE 360

 361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
        | |:||| | |||  :||||| |: |:|:| :||:|| :|||||| : :|| |||||
 361 MGVFEIISKDIGAPLVTIALKDSSKHSVFKIADTIRRFGWTIPAYTMPKDVEHIAVLRVV 420

 421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDID 479
     |||||||:|||||  |::||: |||  |||     ::   |:  | :   | :||| ::
 421 IREDFSRSLAERLANDMKKVLVELDIHPSRT---TTIAHVKAVENGNGNYV-IKKSIVE 475

SULT 14
-10-156-761-12189
Sequence 12189, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
 APPLICANT: OMURA, SATOSHI
 APPLICANT:  IKEDA, HARUO
 APPLICANT:  ISHIKAWA, JUN
 APPLICANT:  HORIKAWA, HIROSHI
 APPLICANT:  SHIBA, TADAYOSHI
 APPLICANT:  SAKAKI, YOSHIYUKI
 APPLICANT:  HATTORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 CURRENT FILING DATE:  2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 12189
  LENGTH: 470
  TYPE: PRT
  ORGANISM: Streptomyces avermitilis
-10-156-761-12189

Query Match             41.4%;  Score 1082;  DB 15;  Length 470;
Best Local Similarity   49.8%;  Pred. No. 4.7e-96;
Matches  209;  Conservative   74;  Mismatches  131;  Indels    6;  Gaps    4;

y          28 PRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDM 87
              |: :: :  :|   |||:::|||||||| |||||:||||||||:   |:    :|| :|
b          35 PKHRLADGPLPPSTAYQLVHDELMLDGNARLNLATFVTTWMEPQAGVLMAECQDKNMIDK 94

y          88 DEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKA 147
              |||| | ||: ||| |:| |:|||  :|  |||  | ||||| |||||| ||:|  :| |
b          95 DEYPRTAELEKRCVAMLADLWNAP--DAGAAVGCSTTGSSEACMLAGLALKRRWA-RRNA 151

y         148 EGKPVD--KPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTIC 205
              :  |    :||:| | |||||||||  ::|||:::| |    | :||  | :: |||||
```

```
Db        152 DRYPARDVRPNLVMGINVQVCWEKFCNFWEVEMRQVPLEGERYHLDPGAAAELCDENTIG 211

Qy        206 VADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRL 265
              |  ||||| :| :| :  |   |    : || | |:||| |||  |||||  :| |||||
Db        212 VVGILGSTFDGSYEPIAELCAALDALQERTGLDIPVHVDGASGAMIAPFLDEDLVWDFRL 271

Qy        266 PLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQV 325
              | | ||| ||||||||  |:|| :||: | |||||:| :|||| : ||| ||||:  :||
Db        272 PRVASINTSGHKYGLVSPGVGWALWRDSEALPEELVFRVNYLGGNMPTFALNFSRPGAQV 331

Qy        326 IAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAF-SLKDSS 384
              :||||  :||| ||:| | :  |     |  | :     |::::: : :|: || :  : :
Db        332 VAQYYTFLRLGREGFRAVQQTTRNVARSLAERVAALGDFHLLTRGDELPVFAFTTAPEVA 391

Qy        385 CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMREL 444
               :  |::|  :| :||:||||| ||| : ::||||| |  ||  |||  | |: ::: :|
Db        392 SYDVFDVSRRMREHGWLVPAYTFPPNREDLSVLRVVCRNGFSTDLAELFVEDLSRLLPDL 451

RESULT 15
US-10-156-761-11138
; Sequence 11138, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11138
;   LENGTH: 454
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-11138

  Query Match             40.2%;  Score 1051;  DB 15;  Length 454;
  Best Local Similarity   45.6%;  Pred. No. 4.5e-93;
  Matches  198;  Conservative   86;  Mismatches  144;  Indels    6;  Gaps    3;

Qy         14 VHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECD 73
              | : | :|::    |    ||  :      ::::::|:::|:|: |||:|||||||||
Db          6 VAALFGNRFLTEPAPSQTFPEEGMTATDTMRLLDEDLVMEGDPQRNLATFVTTWMEPEAQ 65

Qy         74 KLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLA 133
              ::|  ::::|::|  |||:: |:: ||| |:| ||:||        |  | ||||||||
Db         66 RIIAENLHRNFIDHAEYPISAEIEQRCVRMLADLFHAP----GKTTGCRTQGSSEAIMLG 121

Qy        134 GLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQ 193
               |: | ||: :|:|   | |:||:| | :| | |||| |||:|| : | |:|  | : |:
Db        122 ALSLKWKWRERRQAANLPADRPNLVFGGDVHVVWEKFCRYFDVEPRIVPLAEDKYTIGPE 181

Qy        194 QAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAP 253
                   :||||| |  ::|:|  |  :||  :: ||  :  ||   | ||||| |||||: |
Db        182 DVEPHIDENTIGVVAVVGTTFTGHKDDVVGIDKLLRDVRKERDLDIPIHVDGASGGFVWP 241

Qy        254 FLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPT 313
              ||||: :|||||  |:||||||||||||| ||||:::| : || ::|:|: ||||    |
Db        242 FLYPDSKWDFRLEQVRSINVSGHKYGLVYPGIGWLVFREESDLAKDLVFYENYLGKTDAT 301

Qy        314 FTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKD-EG 372
              |||||| |:| |:||||  :||| :||  |||  ::|   | : |  : || ::  | |
Db        302 FTLNFSTGASMVLAQYYNFVRLGRQGYTYVMETMQKNAHALADNLRSSGRFEVIGSDLEQ 361
```

```
373 VPLVAFSLKDSSCHTEFEIS-DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAE 431
    :||||| |     : | :|:  :    ||:|||||:||||: : :|| :::|  ||   |
362 LPLVAFRLAGEHAYDESDIAWQLSAERGWMVPAYTLPPNAERVKILRALVKETLSREQIE 421

432 RLVIDIEKVMRELD 445
    ||  ||      ||
422 RLTQDIADACATLD 435

arch completed: October 22, 2003, 15:24:12
b time : 444 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

M protein - protein search, using sw model

un on:          October 22, 2003, 15:04:24 ; Search time 42 Seconds
                                            (without alignments)
                                            505.715 Million cell updates/sec

itle:           US-10-006-852-2
erfect score:   2615
equence:        1 MVLSHAVSESDVSVHSTFAS..........DIITGWKKFVADRKKTSGIC 502

coring table:   BLOSUM62
                Gapop 10.0 , Gapext 0.5

earched:        328717 seqs, 42310858 residues

otal number of hits satisfying chosen parameters:      328717

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :        Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES
                  %
esult            Query
 No.    Score    Match Length DB  ID                        Description
------------------------------------------------------------------------------
   1     1023     39.1    466  3  US-09-068-195-25          Sequence 25, Appl
   2      325     12.4    589  4  US-08-939-309-8           Sequence 8, Appli
   3      325     12.4    589  4  US-09-849-180-8           Sequence 8, Appli
   4      325     12.4    589  4  US-09-356-643B-2          Sequence 2, Appli
   5      288     11.0    552  4  US-09-356-643B-11         Sequence 11, Appl
   6    264.5     10.1    542  4  US-08-939-309-6           Sequence 6, Appli
   7    264.5     10.1    542  4  US-09-849-180-6           Sequence 6, Appli
   8    264.5     10.1    542  4  US-09-356-643B-4          Sequence 4, Appli
   9    246.5      9.4    568  4  US-08-939-309-2           Sequence 2, Appli
  10    246.5      9.4    568  4  US-09-849-180-2           Sequence 2, Appli
  11    246.5      9.4    568  4  US-09-356-643B-6          Sequence 6, Appli
  12      227      8.7    568  4  US-08-939-309-4           Sequence 4, Appli
  13      227      8.7    568  4  US-09-849-180-4           Sequence 4, Appli
  14      227      8.7    568  4  US-09-356-643B-8          Sequence 8, Appli
  15      224      8.6    568  3  US-09-238-373-2           Sequence 2, Appli
  16      224      8.6    568  4  US-09-740-369-2           Sequence 2, Appli
  17    160.5      6.1    488  4  US-08-939-309-10          Sequence 10, Appl
  18    160.5      6.1    488  4  US-09-849-180-10          Sequence 10, Appl
  19    160.5      6.1    488  4  US-09-356-643B-10         Sequence 10, Appl
  20    133.5      5.1    525  4  US-09-328-352-4492        Sequence 4492, Ap
  21      116      4.4    509  3  US-09-377-557-20          Sequence 20, Appl
  22    107.5      4.1    662  3  US-08-779-814-5           Sequence 5, Appli
  23      105      4.0    729  4  US-09-134-001C-4728       Sequence 4728, Ap
  24    104.5      4.0    784  4  US-09-004-838-12          Sequence 12, Appl
  25    101.5      3.9   1221  4  US-09-107-532A-3959       Sequence 3959, Ap
  26      101      3.9    730  3  US-09-398-865A-2          Sequence 2, Appli
  27      101      3.9    730  4  US-09-710-714-2           Sequence 2, Appli
```

```
  28    100.5      3.8    398  4  US-09-328-352-4507        Sequence 4507, Ap
  29    100.5      3.8    544  3  US-08-687-590-30          Sequence 30, Appl
  30     99.5      3.8    393  3  US-09-377-557-14          Sequence 14, Appl
  31     99.5      3.8    458  4  US-09-134-001C-4422       Sequence 4422, Ap
  32     99.5      3.8    529  3  US-09-464-483-2           Sequence 2, Appli
  33     99.5      3.8    529  3  US-09-414-664-2           Sequence 2, Appli
  34       99      3.8    419  4  US-09-328-352-7844        Sequence 7844, Ap
  35     98.5      3.8   1447  4  US-09-376-330-17          Sequence 17, Appl
  36       98      3.7    623  4  US-09-107-532A-4726       Sequence 4726, Ap
  37     97.5      3.7    402  3  US-09-464-483-4           Sequence 4, Appli
  38     97.5      3.7    402  3  US-09-414-664-4           Sequence 4, Appli
  39       97      3.7    428  3  US-09-052-778-12          Sequence 12, Appl
  40       97      3.7    428  4  US-09-134-001C-4879       Sequence 4879, Ap
  41       97      3.7    990  4  US-09-627-376-7           Sequence 7, Appli
  42       97      3.7   1120  4  US-09-147-404-1           Sequence 1, Appli
  43     96.5      3.7    597  4  US-09-252-991A-32073      Sequence 32073, A
  44     95.5      3.7    391  3  US-09-377-557-18          Sequence 18, Appl
  45     95.5      3.7    508  4  US-09-858-664A-18         Sequence 18, Appl


                                   ALIGNMENTS


RESULT 1
US-09-068-195-25
; Sequence 25, Application US/09068195B
; Patent No. 6140078
; GENERAL INFORMATION:
;  APPLICANT: Sanders, Jan W.
;  APPLICANT: Ledeboer, Adrianus M.
;  APPLICANT: Venema, Gerard
;  APPLICANT: Kok, Jan
;  TITLE OF INVENTION: Salt-Inducible Promoter Derivable from a Lactic Acid
;  TITLE OF INVENTION: Bacterium, and Its Use in a Lacctic Acid Bacterium for
;  TITLE OF INVENTION: Production of Desired Protein
;  FILE REFERENCE: Sanders-60113/0252227
;  CURRENT APPLICATION NUMBER: US/09/068,195B
;  CURRENT FILING DATE: 1998-07-29
;  EARLIER APPLICATION NUMBER: PCT/EP97/04755
;  EARLIER FILING DATE: 1997-08-20
;  EARLIER APPLICATION NUMBER: EP 97200744/7
;  EARLIER FILING DATE: 1997-03-13
;  EARLIER APPLICATION NUMBER: EP 96202444/4
;  EARLIER FILING DATE: 1996-09-05
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 466
;   TYPE: PRT
;   ORGANISM: L. lactis MG1363
US-09-068-195-25

  Query Match             39.1%;  Score 1023;  DB 3;  Length 466;
  Best Local Similarity   45.0%;  Pred. No. 8.3e-95;
  Matches  195;  Conservative   90;  Mismatches  142;  Indels    6;  Gaps    3;

Qy         18 FASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIM 77
              | |   :  ||::|: : ||    |||:: ||:: :|| |||||:|  |:||||  ||:
Db         18 FGSESEQVDLPKYKLAQQSIEPRVAYQLVQDEMLDEGNARLNLATFCQTYMEPEAVKLMS 77

Qy         78 SSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
               :: || :|  ||| |||::|||||||| |:||   |     |  |:||||| || |:|
Db         78 QTLEKNAIDKSEYPRTTEIENRCVNMIADLWNA--SEKGKIYGTSTIGSSEACMLGGMAM 135

Qy        138 KRKWQNKRKAEGKPVD--KPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
              |  |: : :  |  ::  |||:|  :  |||||||  |:::|::|| :   :  :: ::
Db        136 KFSWRKRAEKLGLDINAKKPNLVISSGYQVCWEKFCVYWDIEMREVPMDREHMSINLEKV 195

Qy        196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
              :| ||| || |  |:| |  | ::|:| |::|: | ||:| :   |||||||||  |||:
Db        196 MDYVDEYTIGVVGIMGITYTGRYDDIKALDNLIEEYNKQTDYKVYIHVDAASGGLYAPFV 255

Qy        256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
```

```
                |||||||||  | ||| ||||||||| |:|||:||:|: ||||||| ::||| : ||
           256 EPELEWDFRLKNVISINTSGHKYGLVYPGVGWVLWRDKKYLPEELIFKVSYLGGELPTMA 315

           316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPL 375
               :|||  :||:| |||  :| | :||: : |   :  : | | :|||  | |::    :|:
           316 INFSHSASQLIGQYYNFVRYGFDGYKAIHERTHKVAMYLAEEIEKTGMFEIMNDGAQLPI 375

           376 VAFSLKDSSCH--TEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERL 433
               | : ||::|      ::::| |   || |||| :| | ::  : |:||| ||   :|
           376 VCYKLKENSNRGWNLYDLADRLLMKGWQVPAYPLPKNLENEIIQRLVIRADFGMNMAFNY 435

           434 VIDIEKVMRELDE 446
               | |::: :  |::
           436 VQDMQEAIDALNK 448

SULT 2
-08-939-309-8
 Sequence 8, Application US/08939309
 Patent No. 6423527
  GENERAL INFORMATION:
    APPLICANT:  Saba, Julie D.
    APPLICANT:  Zhou, Jianhui
    TITLE OF INVENTION:  SPHINGOSINE-1-PHOSPHATE LYASE
    TITLE OF INVENTION:  POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
    TITLE OF INVENTION:  METHODS OF USE THEREFOR
    NUMBER OF SEQUENCES:  10
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SEED and BERRY LLP
      STREET:  6300 Columbia Center, 701 Fifth Avenue
      CITY:  Seattle
      STATE:  Washington
      COUNTRY:  USA
      ZIP:  98104
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/939,309
      FILING DATE:  29-SEP-1997
      CLASSIFICATION:  800
    ATTORNEY/AGENT INFORMATION:
      NAME:  David, Maki J.
      REGISTRATION NUMBER:  31,392
      REFERENCE/DOCKET NUMBER:  200116.402
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (206) 622-4900
      TELEFAX:  (206) 682-6031
  INFORMATION FOR SEQ ID NO:  8:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  589 amino acids
      TYPE:  amino acid
      TOPOLOGY:  linear
    MOLECULE TYPE:  protein
-08-939-309-8

Query Match             12.4%;  Score 325;  DB 4;  Length 589;
Best Local Similarity   24.7%;  Pred. No. 6.2e-24;
Matches  122;  Conservative   98;  Mismatches  188;  Indels   86;  Gaps   20;

             3 LSHAVSESDVSVHSTFASRYVR--TSLPRF-KMPENSIPKEAAYQIINDELMLDGNPRLN 59
               |   | :    |  :     :|  : |  | ::| | ||::   : :|         :||
           106 LRGTVEKEVTKVKQSIEDELIRSDSQLMNFPQLPSNGIPQDDVIEELN---------KLN 156

            60 LASFVTTWMEPEC---------DKLIMSSIN------KNYVDMDEYPVTTELQNRCVNMI 104
                    | | :          | : : :|        | :  | :|   ::::  |:|:
           157 DLIPHTQWKEGKVSGAVYHGGDDLIHLQTIAYEKYCVANQLHPDVFPAVRKMESEVVSMV 216

           105 AHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQ 164
                 :||||   ::|  |  | | :|:::|| |: |    : |      : :| |:
           217 LRMFNAP---SDTGCGTTTSGGTESLLLACLSAKMYALHHR-----GITEPEIIAPVTAH 268
```

```
Qy         165 VCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLN---GEFEDV 221
                 ::| | || ::|: |:|    | :|  :    :::|||    ::||  |   |  :|:
Db         269 AGFDKAAYYFGMKLRHVELDPTTYQVDLGKVKKFINKNTIL---LVGSAPNFPHGIADDI 325

Qy         222 KLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL----YPELE-WDFRLPLVKSINVSGH 276
               : |   : :| |      |:|||:  | ||  |:    |  |   |||:| | ||:   |
Db         326 EGLGK-IAQKYK-----LPLHVDSCLGSFIVSFMEKAGYKNLPLLDFRVPGVTSISCDTH 379

Qy         277 KYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLG 336
               |||    |   :::|| :    :   :  : |    : ||  |:  : |:  :  :: :|
Db         380 KYGFAPKGSSVIMYRNSDLRMHQYYVNPAWTGGLYGSPTLAGSRPGAIVVGCWATMVNMG 439

Qy         337 HEGYRNVMENCRENMIVLREGLEKTERFNI----VSKDEGVPLVAFSLKDSSCHTEFEIS 392
                 ||   :|:|:| ::      :|  : ||    :  :    :::|| |  : |   |:|
Db         440 ENGY---IESCQE-IVGAAMKFKKYIQENIPDLNIMGNPRYSVISFSSKTLNIH---ELS 492

Qy         393 DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVI 452
               | | : ||   |   |          | :   |:| |:  :| :|  ::|
Db         493 DRLSKKGWHFNALQKP----------VALHMAFTR-LSAHVVDEICDILR---------- 531

Qy         453 HKISLGQEKSESNS 466
                  :: : ||||||
Db         532 --TTVQELKSESNS 543

RESULT 3
US-09-849-180-8
; Sequence 8, Application US/09849180
; Patent No. 6495359
;   GENERAL INFORMATION:
;        APPLICANT: Saba, Julie D.
;                   Zhou, Jianhui
;        TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
;                            POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
;                            METHODS OF USE THEREFOR
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed Intellectual Property Law Group
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98055
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/849,180
;             FILING DATE: 04-May-2001
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Pepe, Jeffrey C.
;             REGISTRATION NUMBER: 46,985
;             REFERENCE/DOCKET NUMBER: 200116.402
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 589 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-849-180-8

  Query Match             12.4%;  Score 325;  DB 4;  Length 589;
  Best Local Similarity   24.7%;  Pred. No. 6.2e-24;
  Matches  122;  Conservative   98;  Mismatches  188;  Indels   86;  Gaps   20;
```

```
      3 LSHAVSESDVSVHSTFASRYVR--TSLPRF-KMPENSIPKEAAYQIINDELMLDGNPRLN 59
        |   | :    |  :     :|  : |  | ::| | ||::   : :|         :||
    106 LRGTVEKEVTKVKQSIEDELIRSDSQLMNFPQLPSNGIPQDDVIEELN---------KLN 156

     60 LASFVTTWMEPEC---------DKLIMSSIN------KNYVDMDEYPVTTELQNRCVNMI 104
             | | | :          | : : :|        | :  | :|   ::::  |:|:
    157 DLIPHTQWKEGKVSGAVYHGGDDLIHLQTIAYEKYCVANQLHPDVFPAVRKMESEVVSMV 216

    105 AHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQ 164
          :||||   ::|  |  | |:|:::|| |: |    : |      : :| |:
    217 LRMFNAP---SDTGCGTTTSGGTESLLLACLSAKMYALHHR-----GITEPEIIAPVTAH 268

    165 VCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLN---GEFEDV 221
          ::| | || ::|: |:|    | :|  :    :::|||    ::||  |   |  :|:
    269 AGFDKAAYYFGMKLRHVELDPTTYQVDLGKVKKFINKNTIL---LVGSAPNFPHGIADDI 325

    222 KLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL----YPELE-WDFRLPLVKSINVSGH 276
        : |   : :| |      |:|||:  | ||  |:    |  |   |||:| | ||:   |
    326 EGLGK-IAQKYK-----LPLHVDSCLGSFIVSFMEKAGYKNLPLLDFRVPGVTSISCDTH 379

    277 KYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLG 336
        |||    |   :::|| :    :   :  : |    : ||  |:  : |:  :  :: :|
    380 KYGFAPKGSSVIMYRNSDLRMHQYYVNPAWTGGLYGSPTLAGSRPGAIVVGCWATMVNMG 439

    337 HEGYRNVMENCRENMIVLREGLEKTERFNI----VSKDEGVPLVAFSLKDSSCHTEFEIS 392
          ||   :|:|:| ::      :|  : ||    :  :    :::|| |  : |   |:|
    440 ENGY---IESCQE-IVGAAMKFKKYIQENIPDLNIMGNPRYSVISFSSKTLNIH---ELS 492

    393 DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVI 452
        | | : ||   |   |          | :   |:| |:  :| :|  ::|
    493 DRLSKKGWHFNALQKP----------VALHMAFTR-LSAHVVDEICDILR---------- 531

    453 HKISLGQEKSESNS 466
           :: :||||||
    532 --TTVQELKSESNS 543
```

```
SULT 4
-09-356-643B-2
Sequence 2, Application US/09356643B
Patent No. 6569666
GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
 TITLE OF INVENTION:  POLYNUCLEOTIDES AND MODULATING AGENTS AND
 TITLE OF INVENTION:  METHODS OF USE THEREFOR
 FILE REFERENCE: 200116.402C1
 CURRENT APPLICATION NUMBER: US/09/356,643B
 CURRENT FILING DATE:  1999-07-19
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 2
  LENGTH: 589
  TYPE: PRT
  ORGANISM: S. cerevisiae
-09-356-643B-2

Query Match             12.4%;  Score 325;  DB 4;  Length 589;
Best Local Similarity   24.7%;  Pred. No. 6.2e-24;
Matches  122;  Conservative   98;  Mismatches  188;  Indels   86;  Gaps   20;

      3 LSHAVSESDVSVHSTFASRYVR--TSLPRF-KMPENSIPKEAAYQIINDELMLDGNPRLN 59
        |   | :    |  :     :|  : |  | ::| | ||::   : :|         :||
    106 LRGTVEKEVTKVKQSIEDELIRSDSQLMNFPQLPSNGIPQDDVIEELN---------KLN 156

     60 LASFVTTWMEPEC---------DKLIMSSIN------KNYVDMDEYPVTTELQNRCVNMI 104
             | | | :          | : : :|        | :  | :|   ::::  |:|:
    157 DLIPHTQWKEGKVSGAVYHGGDDLIHLQTIAYEKYCVANQLHPDVFPAVRKMESEVVSMV 216

    105 AHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQ 164
          :||||   ::|  |  | |:|:::|| |: |    : |      : :| |:
    217 LRMFNAP---SDTGCGTTTSGGTESLLLACLSAKMYALHHR-----GITEPEIIAPVTAH 268
```

```
Qy        165 VCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLN---GEFEDV 221
                ::| | || ::|: |:|    | :|  :    :::|||    ::||  |   |  :|:
Db        269 AGFDKAAYYFGMKLRHVELDPTTYQVDLGKVKKFINKNTIL---LVGSAPNFPHGIADDI 325

Qy        222 KLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL----YPELE-WDFRLPLVKSINVSGH 276
              : |   : :| |      |:|||:  | ||  |:    |  |   |||:| | ||:   |
Db        326 EGLGK-IAQKYK-----LPLHVDSCLGSFIVSFMEKAGYKNLPLLDFRVPGVTSISCDTH 379

Qy        277 KYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLG 336
              |||    |   :::|| :    :   :  : |    : ||  |:  : |:  :  :: :|
Db        380 KYGFAPKGSSVIMYRNSDLRMHQYYVNPAWTGGLYGSPTLAGSRPGAIVVGCWATMVNMG 439

Qy        337 HEGYRNVMENCRENMIVLREGLEKTERFNI----VSKDEGVPLVAFSLKDSSCHTEFEIS 392
                ||   :|:|:| ::      :|  : ||    :  :    :::|| |  : |   |:|
Db        440 ENGY---IESCQE-IVGAAMKFKKYIQENIPDLNIMGNPRYSVISFSSKTLNIH---ELS 492

Qy        393 DMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVI 452
              | | : ||   |   |          | :   |:| |:  :| :|  ::|
Db        493 DRLSKKGWHFNALQKP----------VALHMAFTR-LSAHVVDEICDILR---------- 531

Qy        453 HKISLGQEKSESNS 466
                 :: :||||||
Db        532 --TTVQELKSESNS 543


RESULT 5
US-09-356-643B-11
; Sequence 11, Application US/09356643B
; Patent No. 6569666
; GENERAL INFORMATION:
;  APPLICANT: Saba, Julie D.
;  TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND MODULATING AGENTS AND
;  TITLE OF INVENTION:  METHODS OF USE THEREFOR
;  FILE REFERENCE: 200116.402C1
;  CURRENT APPLICATION NUMBER: US/09/356,643B
;  CURRENT FILING DATE:  1999-07-19
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
;   LENGTH: 552
;   TYPE: PRT
;   ORGANISM: C. elegans
US-09-356-643B-11

  Query Match             11.0%;  Score 288;  DB 4;  Length 552;
  Best Local Similarity   24.2%;  Pred. No. 3.1e-20;
  Matches  119;  Conservative   89;  Mismatches  208;  Indels   76;  Gaps   21;

Qy         13 SVHSTFASRYVRTSLPRFKMPENSI----PKEAAYQIINDELMLDGNPRLNLASFVTTWM 68
              |:|     :   ::||   : ::||     |   |   |    :||     ::  | |
Db         95 SIHKDDKDKQFISTLPIAPLSQDSIMELAKKYEDYNTFN----IDGG---RVSGAVYTDR 147

Qy         69 EPECDKLIMSSINK----NYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTV 124
                |   |:     |    | :  | :|   :::   : |: :|:| |    | : |  |
Db        148 HAEHINLLGKIYEKYAFSNPLHPDVFPGARKMEAELIRMVLNLYNGP----EDSSGSVTS 203

Qy        125 GSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLS 184
              | :|:|::|  ::      : :|    :: | |:        ::| |    : |: | :
Db        204 GGTESIIMACFSY------RNRAHSLGIEHPVILACKTAHAAFDKAAHLCGMRLRHVPVD 257

Qy        185 EGYYVMDPQQAVDMVDENTICVADILGSTLN---GEFEDVKLLNDLLVEKNKETGWDTPI 241
                  | | ::   ::| | :|:  ::||  |   |  : :  :  |     |: |   |:
Db        258 SDNRV-DLKEMERLIDSN-VCM--LVGSAPNFPSGTIDPIPEIAKL----GKKYG--IPV 307

Qy        242 HVDAASGGFIAPFL----YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLP 297
              ||||  |||: ||:    |    :||| | | ||:   ||||    |   |::|:||
Db        308 HVDACLGGFMIPFMNDAGYLIPVFDFRNPGVTSISCDTHKYGCTPKGSSIVMYRSKE--- 364

Qy        298 EELIFHINYL------GADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENM 351
                 : |  |       |    | |:  |:  :     :  |:  | : |        ::
Db        365 ---LHHFQYFSVADWCGGIYATPTIAGSRAGANTAVAWATLLSFGRDEYVRRCAQIVKHT 421
```

```
y          352 IVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNA 411
                :| | :|| :      | : | |||||    :    :|:|| : : || :      | |
b          422 RMLAEKIEKIKWIKPYGKSD-VSLVAFSGNGVNI---YEVSDKMMKLGWNLNT-LQNPAA 476

y          412 QHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSD---N 468
                || :      |:     |    :|:||:  ||           : |::|::|
b          477 IHICLTINQANEEVVNAFA----VDLEKICEEL----------AAKGEQKADSGMAAMYG 522

y          469 LMVTVKKSDIDK 480
               :   | || :|:
b          523 MAAQVPKSVVDE 534

ESULT 6
S-08-939-309-6
 Sequence 6, Application US/08939309
 Patent No. 6423527
  GENERAL INFORMATION:
    APPLICANT:  Saba, Julie D.
    APPLICANT:  Zhou, Jianhui
    TITLE OF INVENTION:  SPHINGOSINE-1-PHOSPHATE LYASE
    TITLE OF INVENTION:  POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
    TITLE OF INVENTION:  METHODS OF USE THEREFOR
    NUMBER OF SEQUENCES:  10
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SEED and BERRY LLP
      STREET:  6300 Columbia Center, 701 Fifth Avenue
      CITY:  Seattle
      STATE:  Washington
      COUNTRY:  USA
      ZIP:  98104
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/939,309
      FILING DATE:  29-SEP-1997
      CLASSIFICATION:  800
    ATTORNEY/AGENT INFORMATION:
      NAME:  David, Maki J.
      REGISTRATION NUMBER:  31,392
      REFERENCE/DOCKET NUMBER:  200116.402
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (206) 622-4900
      TELEFAX:  (206) 682-6031
  INFORMATION FOR SEQ ID NO:  6:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  542 amino acids
      TYPE:  amino acid
      TOPOLOGY:  linear
    MOLECULE TYPE:  protein
-08-939-309-6

 Query Match             10.1%;  Score 264.5;  DB 4;  Length 542;
 Best Local Similarity   22.3%;  Pred. No. 7.2e-18;
 Matches  109;  Conservative   93;  Mismatches  207;  Indels   79;  Gaps   20;

           3 LSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAA-YQIINDELMLDGNPRLNLA 61
             |:    | : |:     |    |::|   :    : : || |  :     |:|  |:: |
          77 LNEVKDELEKSLRIVDRSTEYFTTIPSHSVGRTEVLRLAAIYDDLEGPAFLEG--RVSGA 134

          62 SF--------------------VTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMI 104
              |                    |  |   ||              :|    ::   | |
         135 VFNREDDKDEREMYEEVFGKFAWTNPLWPKL--------------FPGVRIMEAEVVRMC 180

         105 AHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQ 164
              :: |    ::||  |  : | | :|:|| || :    |:    |:   :  ::  ::|
         181 CNMMNG---DSETC-GTMSTGGSISILLACLAHR----NRLLKRGEKYTE--MIVPSSVH 230

         165 VCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLN---GEFEDV 221
               : | |  | :::::: :    : :|  :    ::: | |:  ::||  |   |  :|:
```

```
Db         231 AAFFKAAECFRIKVRKIPVDPVTFKVDLVKMKAAINKRT-CM--LVGSAPNFPFGTVDDI 287

Qy         222 KLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL-YPELEWDFRLPLVKSINVSGHKYGL 280
               : :  | :|      :| |:||||  |||: |||   |: :|||:| | ||:   |||||
Db         288 EAIGQLGLE------YDIPVHVDACLGGFLLPFLEEDEIRYDFRVPGVSSISADSHKYGL 341

Qy         281 VYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGY 340
                  |   |::|||| |  :     :: |    : |:  |:    :   :  ::    |||
Db         342 APKGSSVVLYRNKELLHNQYFCDADWQGGIYASATMEGSRAGHNIALCWAAMLYHAQEGY 401

Qy         341 ----RNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLR 396
                   | :::  |:    :| ||   :   :    : | :|:::  |      :   : ::
Db         402 KANARKIVDTTRK----IRNGLSNIKGIKLQGPSD-VCIVSWTTNDGV--ELYRFHNFMK 454

Qy         397 RYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKIS 456
                  | :     |    ||    :|        |||  | |    :  :        || |
Db         455 EKHWQLNGLQFPAGV-HI----MVTMNHTHPGLAEAFVADCRAAVEFVKS------HKPS 503

Qy         457 LGQEKSES 464
                  : ||:
Db         504 ESDKTSEA 511


RESULT 7
US-09-849-180-6
; Sequence 6, Application US/09849180
; Patent No. 6495359
;  GENERAL INFORMATION:
;       APPLICANT: Saba, Julie D.
;                  Zhou, Jianhui
;       TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
;                           POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
;                           METHODS OF USE THEREFOR
;       NUMBER OF SEQUENCES: 10
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Seed Intellectual Property Law Group
;            STREET: 701 Fifth Avenue, Suite 6300
;            CITY: Seattle
;            STATE: Washington
;            COUNTRY: USA
;            ZIP: 98055
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/849,180
;            FILING DATE: 04-May-2001
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Pepe, Jeffrey C.
;            REGISTRATION NUMBER: 46,985
;            REFERENCE/DOCKET NUMBER: 200116.402
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (206) 622-4900
;            TELEFAX: (206) 682-6031
;  INFORMATION FOR SEQ ID NO: 6:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 542 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-849-180-6

  Query Match             10.1%;  Score 264.5;  DB 4;  Length 542;
  Best Local Similarity   22.3%;  Pred. No. 7.2e-18;
  Matches  109;  Conservative   93;  Mismatches  207;  Indels   79;  Gaps   20;

Qy           3 LSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAA-YQIINDELMLDGNPRLNLA 61
               |:    | : |:     |    |::|   :    : : || |  :     |:|  |:: |
Db          77 LNEVKDELEKSLRIVDRSTEYFTTIPSHSVGRTEVLRLAAIYDDLEGPAFLEG--RVSGA 134
```

```
          62 SF-----------------VTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMI 104
              |                    |  |   ||              :|    ::   | |
         135 VFNREDDKDEREMYEEVFGKFAWTNPLWPKL--------------FPGVRIMEAEVVRMC 180

         105 AHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQ 164
              :: |    ::||  |  : | | :|:|| || :    |:    |:   :  ::  ::|
         181 CNMMNG---DSETC-GTMSTGGSISILLACLAHR----NRLLKRGEKYTE--MIVPSSVH 230

         165 VCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLN---GEFEDV 221
               : | |  | :::::: :    : :|  :    ::: | |:  ::||  |   |  :|:
         231 AAFFKAAECFRIKVRKIPVDPVTFKVDLVKMKAAINKRT-CM--LVGSAPNFPFGTVDDI 287

         222 KLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL-YPELEWDFRLPLVKSINVSGHKYGL 280
             : :  | :|      :| |:||||  |||: |||   |: :|||:| | ||:   |||||
         288 EAIGQLGLE------YDIPVHVDACLGGFLLPFLEEDEIRYDFRVPGVSSISADSHKYGL 341

         281 VYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGY 340
                |   |::|||| |  :     :: |    : |:  |:    :   :  ::    |||
         342 APKGSSVVLYRNKELLHNQYFCDADWQGGIYASATMEGSRAGHNIALCWAAMLYHAQEGY 401

         341 ----RNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLR 396
                 | :::  |:    :| ||   :   :    : | :|:::  |      :   : ::
         402 KANARKIVDTTRK----IRNGLSNIKGIKLQGPSD-VCIVSWTTNDGV--ELYRFHNFMK 454

         397 RYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKIS 456
                | :     |    ||    :|        |||  | |    :  :        || |
         455 EKHWQLNGLQFPAGV-HI----MVTMNHTHPGLAEAFVADCRAAVEFVKS------HKPS 503

         457 LGQEKSES 464
                : ||:
         504 ESDKTSEA 511
```

```
.SULT 8
-09-356-643B-4
Sequence 4, Application US/09356643B
Patent No. 6569666
GENERAL INFORMATION:
 APPLICANT: Saba, Julie D.
 TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
 TITLE OF INVENTION:  POLYNUCLEOTIDES AND MODULATING AGENTS AND
 TITLE OF INVENTION:  METHODS OF USE THEREFOR
 FILE REFERENCE: 200116.402C1
 CURRENT APPLICATION NUMBER: US/09/356,643B
 CURRENT FILING DATE:  1999-07-19
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 4
  LENGTH: 542
  TYPE: PRT
  ORGANISM: C. elegans
-09-356-643B-4

Query Match             10.1%;  Score 264.5;  DB 4;  Length 542;
Best Local Similarity   22.3%;  Pred. No. 7.2e-18;
Matches  109;  Conservative   93;  Mismatches  207;  Indels   79;  Gaps   20;

           3 LSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAA-YQIINDELMLDGNPRLNLA 61
             |:    | : |:     |    |::|   :    : : || |  :     |:|  |:: |
          77 LNEVKDELEKSLRIVDRSTEYFTTIPSHSVGRTEVLRLAAIYDDLEGPAFLEG--RVSGA 134

          62 SF-----------------VTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMI 104
              |                    |  |   ||              :|    ::   | |
         135 VFNREDDKDEREMYEEVFGKFAWTNPLWPKL--------------FPGVRIMEAEVVRMC 180

         105 AHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQ 164
              :: |    ::||  |  : | | :|:|| || :    |:    |:   :  ::  ::|
         181 CNMMNG---DSETC-GTMSTGGSISILLACLAHR----NRLLKRGEKYTE--MIVPSSVH 230

         165 VCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLN---GEFEDV 221
               : | |  | :::::: :    : :|  :    ::: | |:  ::||  |   |  :|:
         231 AAFFKAAECFRIKVRKIPVDPVTFKVDLVKMKAAINKRT-CM--LVGSAPNFPFGTVDDI 287
```

```
Qy       222 KLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL-YPELEWDFRLPLVKSINVSGHKYGL 280
             : :  | :|      :| |:||||  |||: |||   |: :|||:| | ||:   |||||
Db       288 EAIGQLGLE------YDIPVHVDACLGGFLLPFLEEDEIRYDFRVPGVSSISADSHKYGL 341

Qy       281 VYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGY 340
                |   |::|||| |  :     :: |    : |:  |:    :   :  ::    |||
Db       342 APKGSSVVLYRNKELLHNQYFCDADWQGGIYASATMEGSRAGHNIALCWAAMLYHAQEGY 401

Qy       341 ----RNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLR 396
                 | :::  |:    :| ||   :   :    : | :|:::  |      :   : ::
Db       402 KANARKIVDTTRK----IRNGLSNIKGIKLQGPSD-VCIVSWTTNDGV--ELYRFHNFMK 454

Qy       397 RYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKIS 456
                | :     |    ||    :|        |||  | |    :  :        || |
Db       455 EKHWQLNGLQFPAGV-HI----MVTMNHTHPGLAEAFVADCRAAVEFVKS------HKPS 503

Qy       457 LGQEKSES 464
                : ||:
Db       504 ESDKTSEA 511
```

```
RESULT 9
US-08-939-309-2
; Sequence 2, Application US/08939309
; Patent No. 6423527
;  GENERAL INFORMATION:
;    APPLICANT:  Saba, Julie D.
;    APPLICANT:  Zhou, Jianhui
;    TITLE OF INVENTION:  SPHINGOSINE-1-PHOSPHATE LYASE
;    TITLE OF INVENTION:  POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
;    TITLE OF INVENTION:  METHODS OF USE THEREFOR
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/939,309
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  David, Maki J.
;      REGISTRATION NUMBER:  31,392
;      REFERENCE/DOCKET NUMBER:  200116.402
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  568 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-939-309-2

  Query Match             9.4%;  Score 246.5;  DB 4;  Length 568;
  Best Local Similarity   23.2%;  Pred. No. 5.2e-16;
  Matches  100;  Conservative   76;  Mismatches  210;  Indels   45;  Gaps   12;

Qy        22 YVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSIN 81
             ||:|      :|   :      : : :   :||: :   ||      ||:  :|:: :
Db       112 YVKT------LPAQGMGTAEVLERLKEYSSMDGSWQEGKASGAVYNGEPKLTELLVQAYG 165

Qy        82 K----NYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
```

```
              :    | :  | :|   :|:   | |   |||        : |  | | :|:|::|  |:
b        166 EFTWSNPLHPDIFPGLRKLEAEIVRMTCSLFNG----GPDSCGCVTSGGTESILMACKAY 221

y        138 KRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVD 197
                   :  |  | :  | ||   :    ::| | || :::  | | :   | | |
b        222 ------RDLALEKGIKTPEIVAPESAHAAFDKAAHYFGMKIVRVALKKNMEV-DVQAMKR 274

y        198 MVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL-- 255
              :  ||  :        :|  : |  :  | |       :  |:||||  |||:  |:
b        275 AISRNTAMLVCSTPQFPHGVMDPVPEVAKLTVR------YKIPLHVDACLGGFLIVFMEK 328

y        256 --YP-ELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQP 312
               || |  :|||:  | ||:   ||||    |   |:: |::    :     :: |
b        329 AGYPLEKPFDFRVKGVTSISADTHKYGYAPKGSSVVMYSNEKYRTYQFFVGADWQGGVYA 388

y        313 TFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEG 372
             : ::  |:    : | :  |:  |  ||    :   :    |:  ||  :   |   |
b        389 SPSIAGSRPGGIIAACWAALMHFGENGYVEATKQIIKTARFLKSELENIKNIFIFG-DPQ 447

y        373 VPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAER 432
             : ::|    |      : :|:|:   ||    |   | : |  :  |  |        :|
b        448 LSVIALGSNDFDI---YRLSNMMSAKGWNF-NYLQFPRSIHFCITLVHTR--------KR 495

y        433 LVIDIEKVMRE 443
             : |   | :||
b        496 VAIQFLKDIRE 506

ESULT 10
S-09-849-180-2
 Sequence 2, Application US/09849180
 Patent No. 6495359
   GENERAL INFORMATION:
        APPLICANT: Saba, Julie D.
                   Zhou, Jianhui
        TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
                            POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
                            METHODS OF USE THEREFOR
        NUMBER OF SEQUENCES: 10
        CORRESPONDENCE ADDRESS:
             ADDRESSEE: Seed Intellectual Property Law Group
             STREET: 701 Fifth Avenue, Suite 6300
             CITY: Seattle
             STATE: Washington
             COUNTRY: USA
             ZIP: 98055
        COMPUTER READABLE FORM:
             MEDIUM TYPE: Floppy disk
             COMPUTER: IBM PC compatible
             OPERATING SYSTEM: PC-DOS/MS-DOS
             SOFTWARE: PatentIn Release #1.0, Version #1.30
        CURRENT APPLICATION DATA:
             APPLICATION NUMBER: US/09/849,180
             FILING DATE: 04-May-2001
             CLASSIFICATION: <Unknown>
        ATTORNEY/AGENT INFORMATION:
             NAME: Pepe, Jeffrey C.
             REGISTRATION NUMBER: 46,985
             REFERENCE/DOCKET NUMBER: 200116.402
        TELECOMMUNICATION INFORMATION:
             TELEPHONE: (206) 622-4900
             TELEFAX: (206) 682-6031
    INFORMATION FOR SEQ ID NO: 2:
        SEQUENCE CHARACTERISTICS:
             LENGTH: 568 amino acids
             TYPE: amino acid
             TOPOLOGY: linear
        MOLECULE TYPE: protein
        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 -09-849-180-2

 Query Match             9.4%;  Score 246.5;  DB 4;  Length 568;
 Best Local Similarity   23.2%;  Pred. No. 5.2e-16;
```

```
  Matches  100;  Conservative   76;  Mismatches  210;  Indels   45;  Gaps   12;

Qy         22 YVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSIN 81
              ||:|      :|   :      : : :   :||: :   ||      ||:  :|:: :
Db        112 YVKT------LPAQGMGTAEVLERLKEYSSMDGSWQEGKASGAVYNGEPKLTELLVQAYG 165

Qy         82 K----NYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
              :    | :  | :|   :|:   | |   |||        : |  | | :|:|::|  |:
Db        166 EFTWSNPLHPDIFPGLRKLEAEIVRMTCSLFNG----GPDSCGCVTSGGTESILMACKAY 221

Qy        138 KRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVD 197
                    :  |  | :  | ||   :    ::| | || :::  | | :   | | |
Db        222 ------RDLALEKGIKTPEIVAPESAHAAFDKAAHYFGMKIVRVALKKNMEV-DVQAMKR 274

Qy        198 MVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL-- 255
               :  ||  :        :|  : |  :  | |       :  |:||||  |||:  |:
Db        275 AISRNTAMLVCSTPQFPHGVMDPVPEVAKLTVR------YKIPLHVDACLGGFLIVFMEK 328

Qy        256 --YP-ELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQP 312
                || |  :|||:  | ||:   ||||    |   |:: |::    :     :: |
Db        329 AGYPLEKPFDFRVKGVTSISADTHKYGYAPKGSSVVMYSNEKYRTYQFFVGADWQGGVYA 388

Qy        313 TFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEG 372
              : ::  |:    : | :  |:  |  ||    :   :    |:  ||  :   |   |
Db        389 SPSIAGSRPGGIIAACWAALMHFGENGYVEATKQIIKTARFLKSELENIKNIFIFG-DPQ 447

Qy        373 VPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAER 432
              : ::|    |      : :|:|:   ||    |   | : |  :  |  |        :|
Db        448 LSVIALGSNDFDI---YRLSNMMSAKGWNF-NYLQFPRSIHFCITLVHTR--------KR 495

Qy        433 LVIDIEKVMRE 443
              : |   | :||
Db        496 VAIQFLKDIRE 506

RESULT 11
US-09-356-643B-6
; Sequence 6, Application US/09356643B
; Patent No. 6569666
; GENERAL INFORMATION:
;  APPLICANT: Saba, Julie D.
;  TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND MODULATING AGENTS AND
;  TITLE OF INVENTION:  METHODS OF USE THEREFOR
;  FILE REFERENCE: 200116.402C1
;  CURRENT APPLICATION NUMBER: US/09/356,643B
;  CURRENT FILING DATE:  1999-07-19
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 568
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-356-643B-6

  Query Match             9.4%;  Score 246.5;  DB 4;  Length 568;
  Best Local Similarity   23.2%;  Pred. No. 5.2e-16;
  Matches  100;  Conservative   76;  Mismatches  210;  Indels   45;  Gaps   12;

Qy         22 YVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSIN 81
              ||:|      :|   :      : : :   :||: :   ||      ||:  :|:: :
Db        112 YVKT------LPAQGMGTAEVLERLKEYSSMDGSWQEGKASGAVYNGEPKLTELLVQAYG 165

Qy         82 K----NYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
              :    | :  | :|   :|:   | |   |||        : |  | | :|:|::|  |:
Db        166 EFTWSNPLHPDIFPGLRKLEAEIVRMTCSLFNG----GPDSCGCVTSGGTESILMACKAY 221

Qy        138 KRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVD 197
                    :  |  | :  | ||   :    ::| | || :::  | | :   | | |
Db        222 ------RDLALEKGIKTPEIVAPESAHAAFDKAAHYFGMKIVRVALKKNMEV-DVQAMKR 274

Qy        198 MVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL-- 255
               :  ||  :        :|  : |  :  | |       :  |:||||  |||:  |:
```

```
275 AISRNTAMLVCSTPQFPHGVMDPVPEVAKLTVR------YKIPLHVDACLGGFLIVFMEK 328

256 --YP-ELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQP 312
      || |  :|||:  | ||:   ||||    |   |:: |::    :     :: |
329 AGYPLEKPFDFRVKGVTSISADTHKYGYAPKGSSVVMYSNEKYRTYQFFVGADWQGGVYA 388

313 TFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEG 372
    : ::  |:    : | :  |:  |  ||    :   :    |:  ||  :   |   |
389 SPSIAGSRPGGIIAACWAALMHFGENGYVEATKQIIKTARFLKSELENIKNIFIFG-DPQ 447

373 VPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAER 432
    : ::|    |      : :|:|:   ||    |   | : |  :  |  |        :|
448 LSVIALGSNDFDI---YRLSNMMSAKGWNF-NYLQFPRSIHFCITLVHTR--------KR 495

433 LVIDIEKVMRE 443
    : |   | :||
496 VAIQFLKDIRE 506

SULT 12
-08-939-309-4
Sequence 4, Application US/08939309
Patent No. 6423527
 GENERAL INFORMATION:
   APPLICANT:  Saba, Julie D.
   APPLICANT:  Zhou, Jianhui
   TITLE OF INVENTION:  SPHINGOSINE-1-PHOSPHATE LYASE
   TITLE OF INVENTION:  POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
   TITLE OF INVENTION:  METHODS OF USE THEREFOR
   NUMBER OF SEQUENCES:  10
   CORRESPONDENCE ADDRESS:
     ADDRESSEE:  SEED and BERRY LLP
     STREET:  6300 Columbia Center, 701 Fifth Avenue
     CITY:  Seattle
     STATE:  Washington
     COUNTRY:  USA
     ZIP:  98104
   COMPUTER READABLE FORM:
     MEDIUM TYPE:  Floppy disk
     COMPUTER:  IBM PC compatible
     OPERATING SYSTEM:  PC-DOS/MS-DOS
     SOFTWARE:  PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER:  US/08/939,309
     FILING DATE:  29-SEP-1997
     CLASSIFICATION:  800
   ATTORNEY/AGENT INFORMATION:
     NAME:  David, Maki J.
     REGISTRATION NUMBER:  31,392
     REFERENCE/DOCKET NUMBER:  200116.402
   TELECOMMUNICATION INFORMATION:
     TELEPHONE:  (206) 622-4900
     TELEFAX:  (206) 682-6031
 INFORMATION FOR SEQ ID NO:  4:
   SEQUENCE CHARACTERISTICS:
     LENGTH:  568 amino acids
     TYPE:  amino acid
     TOPOLOGY:  linear
   MOLECULE TYPE:  protein
-08-939-309-4

Query Match             8.7%;  Score 227;  DB 4;  Length 568;
Best Local Similarity   21.6%;  Pred. No. 4.9e-14;
Matches   98;  Conservative   84;  Mismatches  212;  Indels   60;  Gaps   14;

  8 SESDVSVHSTFA---SRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFV 64
    :: |:| : :|      ||:       :|   :   |  : : :   :|   :   ||
 95 TKDDISKNMSFLKVDKEYVKA------LPSQGLSSSAVLEKLKEYSSMDAFWQEGRASGT 148

 65 TTWMEPECDKLIMSSIN----KNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
        | :  :|:: :       | :  | :|   :::   | :   |||        : |
149 VYSGEEKLTELLVKAYGDFAWSNPLHPDIFPGLRKIEAEIVRIACSLFNG----GPDSCG 204
```

```
Qy      121 VGTVGSSEAIMLA-----GLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFE 175
              | | :|:|::|      |||:           | :  | ||   :    : | | ||
Db      205 CVTSGGTESILMACKACRDLAFE-----------KGIKTPEIVAPQSAHAAFNKAASYFG 253

Qy      176 VELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKET 235
            :::  | |::   | | :     :  ||  :        :|  : |  :  | |:
Db      254 MKIVRVPLTKMMEV-DVRAMRRAISRNTAMLVCSTPQFPHGVIDPVPEVAKLAVK----- 307

Qy      236 GWDTPIHVDAASGGFIAPFL----YP-ELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIW 290
             :  |:||||  |||:  |:    || |  :|||:  | ||:   ||||    |   |::
Db      308 -YKIPLHVDACLGGFLIVFMEKAGYPLEHPFDFRVKGVTSISADTHKYGYAPKGSSLVLY 366

Qy      291 RNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCREN 350
             :|:    :     :: |    : |:  |:      | :  |:  |  ||    :   :
Db      367 SDKKYRNYQFFVDTDWQGGIYASPTIAGSRPGGISAACWAALMHFGENGYVEATKQIIKT 426

Qy      351 MIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPN 410
               |:  ||  :   :    : : |:|   :|      : :|:::   || :     ||:
Db      427 ARFLKSELENIKGIFVFGNPQ-LSLIALGSRDFDI---YRLSNLMTAKGWNLNQLQFPPS 482

Qy      411 AQH-ITVLRVVIREDFSRTLAERLVIDIEKVMRE 443
                ||:|             :|: |   | :||
Db      483 IHFCITLLHA----------RKRVAIQFLKDIRE 506

RESULT 13
US-09-849-180-4
; Sequence 4, Application US/09849180
; Patent No. 6495359
;   GENERAL INFORMATION:
;        APPLICANT: Saba, Julie D.
;                   Zhou, Jianhui
;        TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE
;                            POLYPEPTIDES, POLYNUCLEOTIDES AND MODULATING AGENTS AND
;                            METHODS OF USE THEREFOR
;        NUMBER OF SEQUENCES: 10
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed Intellectual Property Law Group
;             STREET: 701 Fifth Avenue, Suite 6300
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: USA
;             ZIP: 98055
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/849,180
;             FILING DATE: 04-May-2001
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Pepe, Jeffrey C.
;             REGISTRATION NUMBER: 46,985
;             REFERENCE/DOCKET NUMBER: 200116.402
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (206) 622-4900
;             TELEFAX: (206) 682-6031
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 568 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-849-180-4

  Query Match             8.7%;  Score 227;  DB 4;  Length 568;
  Best Local Similarity   21.6%;  Pred. No. 4.9e-14;
  Matches   98;  Conservative   84;  Mismatches  212;  Indels   60;  Gaps   14;

Qy        8 SESDVSVHSTFA---SRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFV 64
```

```
               :: |:| : :|      ||:       :|   :   |  : : :   :|   :   ||
b           95 TKDDISKNMSFLKVDKEYVKA------LPSQGLSSSAVLEKLKEYSSMDAFWQEGRASGT 148

y           65 TTWMEPECDKLIMSSIN----KNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
                   | :  :|:: :       | :  | :|   :::   | :   |||        : |
b          149 VYSGEEKLTELLVKAYGDFAWSNPLHPDIFPGLRKIEAEIVRIACSLFNG----GPDSCG 204

y          121 VGTVGSSEAIMLA-----GLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFE 175
                 | | :|:|::|      |||:           | :  | ||   :    : | | ||
b          205 CVTSGGTESILMACKACRDLAFE-----------KGIKTPEIVAPQSAHAAFNKAASYFG 253

y          176 VELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKET 235
               ::: | |::   | | :     :  ||  :        :|  : |  :  | |:
b          254 MKIVRVPLTKMMEV-DVRAMRRAISRNTAMLVCSTPQFPHGVIDPVPEVAKLAVK----- 307

y          236 GWDTPIHVDAASGGFIAPFL----YP-ELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIW 290
                :  |:||||  |||:  |:    || |  :|||:  | ||:   ||||    |   |::
b          308 -YKIPLHVDACLGGFLIVFMEKAGYPLEHPFDFRVKGVTSISADTHKYGYAPKGSSLVLY 366

y          291 RNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCREN 350
                :|:    :     :: |    : |:  |:      | :  |:  |  ||    :   :
b          367 SDKKYRNYQFFVDTDWQGGIYASPTIAGSRPGGISAACWAALMHFGENGYVEATKQIIKT 426

y          351 MIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPN 410
                  |:  ||  :   :    : |:|   :|      : :|:::   || :     ||:
o          427 ARFLKSELENIKGIFVFGNPQ-LSLIALGSRDFDI---YRLSNLMTAKGWNLNQLQFPPS 482

y          411 AQH-ITVLRVVIREDFSRTLAERLVIDIEKVMRE 443
                   ||:|             :|: |   | :||
o          483 IHFCITLLHA----------RKRVAIQFLKDIRE 506
```

```
ESULT 14
S-09-356-643B-8
 Sequence 8, Application US/09356643B
 Patent No. 6569666
 GENERAL INFORMATION:
  APPLICANT: Saba, Julie D.
  TITLE OF INVENTION: SPHINGOSINE-1-PHOSPHATE LYASE POLYPEPTIDES,
  TITLE OF INVENTION:  POLYNUCLEOTIDES AND MODULATING AGENTS AND
  TITLE OF INVENTION:  METHODS OF USE THEREFOR
  FILE REFERENCE: 200116.402C1
  CURRENT APPLICATION NUMBER: US/09/356,643B
  CURRENT FILING DATE:  1999-07-19
  NUMBER OF SEQ ID NOS: 14
  SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 8
   LENGTH: 568
   TYPE: PRT
   ORGANISM: Homo sapiens
 -09-356-643B-8

 Query Match             8.7%;  Score 227;  DB 4;  Length 568;
 Best Local Similarity   21.6%;  Pred. No. 4.9e-14;
 Matches   98;  Conservative   84;  Mismatches  212;  Indels   60;  Gaps   14;

             8 SESDVSVHSTFA---SRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFV 64
               :: |:| : :|      ||:       :|   :   |  : : :   :|   :   ||
            95 TKDDISKNMSFLKVDKEYVKA------LPSQGLSSSAVLEKLKEYSSMDAFWQEGRASGT 148

            65 TTWMEPECDKLIMSSIN----KNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
                   | :  :|:: :       | :  | :|   :::   | :   |||        : |
           149 VYSGEEKLTELLVKAYGDFAWSNPLHPDIFPGLRKIEAEIVRIACSLFNG----GPDSCG 204

           121 VGTVGSSEAIMLA-----GLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFE 175
                 | | :|:|::|      |||:           | :  | ||   :    : | | ||
           205 CVTSGGTESILMACKACRDLAFE-----------KGIKTPEIVAPQSAHAAFNKAASYFG 253

           176 VELKEVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKET 235
               ::: | |::   | | :     :  ||  :        :|  : |  :  | |:
           254 MKIVRVPLTKMMEV-DVRAMRRAISRNTAMLVCSTPQFPHGVIDPVPEVAKLAVK----- 307

           236 GWDTPIHVDAASGGFIAPFL----YP-ELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIW 290
```

```
                :  |:||||  |||:  |:    || |  :|||:  | ||:   ||||    |   |::
Db         308 -YKIPLHVDACLGGFLIVFMEKAGYPLEHPFDFRVKGVTSISADTHKYGYAPKGSSLVLY 366

Qy         291 RNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCREN 350
                :|:    :     :: |    : |:  |:      | :  |:  |  ||    :   :
Db         367 SDKKYRNYQFFVDTDWQGGIYASPTIAGSRPGGISAACWAALMHFGENGYVEATKQIIKT 426

Qy         351 MIVLREGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPN 410
                  |:  ||  :   :    : |:|   :|      : :|:::   || :     ||:
Db         427 ARFLKSELENIKGIFVFGNPQ-LSLIALGSRDFDI---YRLSNLMTAKGWNLNQLQFPPS 482

Qy         411 AQH-ITVLRVVIREDFSRTLAERLVIDIEKVMRE 443
                   ||:|             :|: |   | :||
Db         483 IHFCITLLHA----------RKRVAIQFLKDIRE 506
```

```
RESULT 15
US-09-238-373-2
; Sequence 2, Application US/09238373A
; Patent No. 6187562
; GENERAL INFORMATION:
;  APPLICANT: DUCKWORTH, DAVID MALCOLM
;  APPLICANT: GODDEN, ROBERT JAMES
;  APPLICANT: TESTA, TANIA TAMSON
;  TITLE OF INVENTION: NOVEL COMPOUNDS
;  FILE REFERENCE: GP-30034
;  CURRENT APPLICATION NUMBER: US/09/238,373A
;  CURRENT FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: UK 9824026.0
;  EARLIER FILING DATE: 1998-11-03
;  EARLIER APPLICATION NUMBER: EP 98300625.5
;  EARLIER FILING DATE: 1998-01-29
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 568
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-238-373-2

  Query Match             8.6%;  Score 224;  DB 3;  Length 568;
  Best Local Similarity   21.4%;  Pred. No. 9.9e-14;
  Matches   96;  Conservative   86;  Mismatches  217;  Indels   50;  Gaps   13;

Qy           8 SESDVSVHSTFA---SRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFV 64
               :: |:| : :|      ||:       :|   :   |  : : :   :|   :   ||
Db          95 TKDDISKNMSFLKVDKEYVKA------LPSQGLSSSAVLEKLKEYSSMDAFWQEGRASGT 148

Qy          65 TTWMEPECDKLIMSSIN----KNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
                   | :  :|:: :       | :  | :|   :::   | :   |||        : |
Db         149 VYSGEEKLTELLVKAYGDFAWSNPLHPDIFPGLRKIEAEIVRIACSLFNG----GPDSCG 204

Qy         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
                 | | :|:|::|  |:      :  |  | :  | ||   :    : | | || :::
Db         205 CVTSGGTESILMACKAY------RDLAFEKGIKTPEIVAPQSAHAAFNKAASYFGMKIVR 258

Qy         181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
               | |::   | | :     :  ||  :        :|  : |  :  | |:      :  |
Db         259 VPLTKMMEV-DVRAMRRAISRNTAMLVCSTPQFPHGVIDPVPEVAKLAVK-----YKIP 311

Qy         241 IHVDAASGGFIAPFL----YP-ELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKED 295
               :||||  |||:  |:    || |  :|||:  | ||:   ||||    |   |:: :|:
Db         312 LHVDACLGGFLIVFMEKAGYPLEHPFDFRVKGVTSISADTHKYGYAPKGSSLVLYSDKKY 371

Qy         296 LPEELIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLR 355
                  :     :: |    : |:  |:      | :  |:  |  ||    :   :    |:
Db         372 RNYQFFVDTDWQGGIYASPTIAGSRPGGISAACWAALMHFGENGYVEATKQIIKTARFLK 431

Qy         356 EGLEKTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQH-I 414
                 ||  :   :    : : ::|   :|      : :|:::   || :     ||:    |
Db         432 SELENIKGIFVFGNPQ-LSVIALGSRDFDI---YRLSNLMTAKGWNLNQLQFPPSIHFCI 487

Qy         415 TVLRVVIREDFSRTLAERLVIDIEKVMRE 443
```

```
            |:|             :|: |   | :||
        488 TLLHA----------RKRVAIQFLKDIRE 506

arch completed: October 22, 2003, 15:09:32
b time : 43 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2003  Compugen Ltd.

1 protein - protein search, using sw model

n on:          October 22, 2003, 13:34:43 ; Search time 69 Seconds
                                           (without alignments)
                                           1154.793 Million cell updates/sec

tle:           US-10-006-852-2
rfect score:   2615
quence:        1 MVLSHAVSESDVSVHSTFAS.........DIITGWKKFVADRKKTSGIC 502

oring table:   BLOSUM62
               Gapop 10.0 , Gapext 0.5

arched:        1107863 seqs, 158726573 residues

tal number of hits satisfying chosen parameters:      1107863

nimum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

tabase :        A_Geneseq_19Jun03:*
               1:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
               2:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
               3:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
               4:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
               5:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
               6:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
               7:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
               8:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
               9:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
               10:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
               11:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
               12:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
               13:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
               14:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
               15:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
               16:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
               17:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
               18:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
               19:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
               20:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
               21:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
               22:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
               23:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
               24:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

  Pred. No. is the number of results predicted by chance to have a
  score greater than or equal to the score of the result being printed,
  and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
esult           Query
 No.    Score   Match Length DB  ID                      Description
------------------------------------------------------------------------
   1     2615   100.0    502  23  AAO15132                A thaliana GAD1.
   2     2607    99.7    502  23  ABB93521                Herbicidally activ
   3     2250    86.0    500  23  AAO15139                Petunia GAD protei
   4     2240    85.7    496  23  AAO15137                Tobacco GAD1.  Nic
   5     2199    84.1    496  23  AAO15138                Tobacco GAD2.  Nic
   6   2135.5    81.7    493  23  AAO15135                A thaliana GAD4.
   7   2135.5    81.7    493  23  ABB91695                Herbicidally activ
   8     2130    81.5    494  23  AAO15133                A thaliana GAD2.
   9     2130    81.5    494  23  ABB91465                Herbicidally activ
```

```
     10     2080    79.5     500  23  AAO15134                A thaliana GAD3.
     11     2080    79.5     500  23  ABB91694                Herbicidally activ
     12     2014    77.0     502  23  AAO15140                Tomato GAD protein
     13     1969    75.3     494  23  AAO15136                A thaliana GAD5.
     14     1969    75.3     494  23  ABB92352                Herbicidally activ
     15     1965    75.1     494  21  AAG13764                Arabidopsis thalia
     16     1965    75.1     507  21  AAG13763                Arabidopsis thalia
     17   1876.5    71.8     464  21  AAG13765                Arabidopsis thalia
     18     1318    50.4     514  22  AAB96992                A oryzae glutamate
     19   1188.5    45.4     479  21  AAG36333                Arabidopsis thalia
     20     1175    44.9     464  21  AAG36334                Arabidopsis thalia
     21   1050.5    40.2     430  21  AAG36335                Arabidopsis thalia
     22     1028    39.3     466  23  ABB54627                Lactococcus lactis
     23     1023    39.1     466  19  AAW59165                L. lactis MG1316 g
     24    955.5    36.5     462  23  ABB47826                Listeria monocytog
     25    953.5    36.5     467  23  ABB48689                Listeria monocytog
     26      940    35.9     464  23  ABB49717                Listeria monocytog
     27      697    26.7    1042  22  ABG17675                Novel human diagno
     28      697    26.7    1042  22  ABG25740                Novel human diagno
     29    626.5    24.0     194  23  ABB92351                Herbicidally activ
     30    508.5    19.4     609  22  ABG17670                Novel human diagno
     31      504    19.3    1155  22  ABG25739                Novel human diagno
     32      504    19.3    1231  22  ABG17676                Novel human diagno
     33      504    19.3    1231  22  ABG24397                Novel human diagno
     34    473.5    18.1     572  22  ABG25741                Novel human diagno
     35    393.5    15.0     191  22  AAU87540                Novel central nerv
     36    393.5    15.0     191  22  AAU22935                Novel human enzyme
     37    393.5    15.0     191  22  AAU18706                Renal and cardiova
     38      343    13.1     213  22  ABG08876                Novel human diagno
     39      325    12.4     589  20  AAY05829                Yeast sphingosine-
     40      325    12.4     589  22  AAG70706                S cerevisiae apopt
     41      314    12.0     384  22  AAB96568                Putative P. abyssi
     42    304.5    11.6     589  22  AAG70849                C albicans apoptos
     43    299.5    11.5     545  22  ABB64094                Drosophila melanog
     44    299.5    11.5     545  22  ABB64099                Drosophila melanog
     45    299.5    11.5     545  22  AAE03543                Drosophila melanog


                              ALIGNMENTS


RESULT 1
AAO15132
ID   AAO15132 standard; Protein; 502 AA.
XX
AC   AAO15132;
XX
DT   25-SEP-2002  (first entry)
XX
DE   A thaliana GAD1.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   N-PSDB; AAL43410.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
```

```
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T    -
X
S   Claim 17; Page 53; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
C   (GABA) in response to a signal, by incorporating into the plant's genome
C   a DNA construct with a non-constitutive promoter operably linked to a
C   polynucleotide encoding a functional plant glutamic acid decarboxylase
C   (GAD), to provide a transformed plant that expresses the GAD coding
C   sequence in response to a signal. Plants of this type have an enhanced
C   ability to tolerate environmental or other stresses. The present sequence
C   is the A. thaliana GAD1 protein.
X
Q   Sequence   502 AA;

 Query Match            100.0%;  Score 2615;  DB 23;  Length 502;
 Best Local Similarity  100.0%;  Pred. No. 1.4e-241;
 Matches  502;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y         1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b         1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

y        61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b        61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120

y       121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180

y       181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240

y       241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300

y       301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360

y       361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420

y       421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b       421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480

y       481 QRDIITGWKKFVADRKKTSGIC 502
            ||||||||||||||||||||||
b       481 QRDIITGWKKFVADRKKTSGIC 502


SULT 2
B93521
)   ABB93521 standard; Protein; 502 AA.
X
:   ABB93521;
X
T   31-MAY-2002  (first entry)
X
E   Herbicidally active polypeptide SEQ ID NO 2732.
X
N   Herbicidal; plant; agriculture; herbicide.
X
S   Arabidopsis thaliana.
X
N   WO200210210-A2.
```

```
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP09892.
XX
PR   28-AUG-2001; 2001WO-EP09892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms -
XX
PS   Claim 5; SEQ ID NO 2732; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins
CC   (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC   aligning and comparing nucleic acid or amino acid sequences from plant
CC   with nucleic acid or amino acid sequences from non-plant organisms using
CC   suitable search parameters, where plant sequences having an E-value
CC   greater by a factor of 3 than the E-value of most similar non-plant
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   502 AA;

  Query Match            99.7%;  Score 2607;  DB 23;  Length 502;
  Best Local Similarity  99.8%;  Pred. No. 7.9e-241;
  Matches  501;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        181 VKLSEGYYVMDPQQAVDMVDENTICVAAILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480

Qy        481 QRDIITGWKKFVADRKKTSGIC 502
              ||||||||||||||||||||||
Db        481 QRDIITGWKKFVADRKKTSGIC 502
```

```
ESULT 3
AO15139
D   AAO15139 standard; Protein; 500 AA.
X
C   AAO15139;
X
T   25-SEP-2002  (first entry)
X
E   Petunia GAD protein.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance.
X
S   Petunia sp.
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I   Kinnersley AM,  Turano FJ;
X
R   WPI; 2002-490073/52.
R   N-PSDB; AAL43417.
X
T   Making transformed plants that selectively increase gamma-aminobutyric
T   acid production, by incorporating a DNA construct with a polynucleotide
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T    -
X
S   Claim 17; Page 58; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
C   (GABA) in response to a signal, by incorporating into the plant's genome
C   a DNA construct with a non-constitutive promoter operably linked to a
C   polynucleotide encoding a functional plant glutamic acid decarboxylase
C   (GAD), to provide a transformed plant that expresses the GAD coding
C   sequence in response to a signal. Plants of this type have an enhanced
C   ability to tolerate environmental or other stresses. The present sequence
C   is the petunia GAD protein.
X
Q   Sequence   500 AA;

 Query Match             86.0%;  Score 2250;  DB 23;  Length 500;
 Best Local Similarity   84.9%;  Pred. No. 1.4e-206;
 Matches  427;  Conservative   39;  Mismatches   33;  Indels    4;  Gaps    2;

y          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
             ||||  ||:||||:|||||||||||||||||||:||||||||||||||||||||||||||
b          1 MVLSKTVSQSDVSIHSTFASRYVRTSLPRFKMPDNSIPKEAAYQIINDELMLDGNPRLNL 60

y         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
             |||||||||||||||:| |||||||||||||||||||||||||||||||||||: |||||
b         61 ASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEDGETAVG 120

y        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
             |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
b        121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTGANVQVCWEKFARYFEVELKE 180

y        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
             ||||||||||||::||:|||||||||| |||||||||||||| |||||||||||||||||
b        181 VKLSEGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVEKNKETGWDTP 240

y        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
             ||||||||||||||:|||||||||||||||||||||||||||||||||:||||:|||:||
b        241 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRNKDDLPDEL 300
```

```
Qy       301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
             ||||||||||||||||||||||||||||||||||||:|||:||||||:||  ||||||||
Db       301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGYEGYKNVMENCQENASVLREGLEK 360

Qy       361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
             | ||||:||: |||||||||||:  | |||||: |||:||||||||||||||||||||||
Db       361 TGRFNIISKEIGVPLVAFSLKDNRQHNEFEISETLRRFGWIVPAYTMPPNAQHITVLRVV 420

Qy       421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
             |||||||||||||| |||||: ||| ||:||  |::: :|:: :|   :    ||:| :
Db       421 IREDFSRTLAERLVRDIEKVLHELDTLPARVNAKLAVAEEQAAANGSEVH---KKTDSEV 477

Qy       481 QRDIITGWKKFVAD-RKKTSGIC 502
             | ::|| ||||| : :|||: :|
Db       478 QLEMITAWKKFVEEKKKKTNRVC 500


RESULT 4
AAO15137
ID   AAO15137 standard; Protein; 496 AA.
XX
AC   AAO15137;
XX
DT   25-SEP-2002  (first entry)
XX
DE   Tobacco GAD1.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance.
XX
OS   Nicotiana tabacum.
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   N-PSDB; AAL43415.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
PT   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT    -
XX
PS   Claim 17; Page 57; 63pp; English.
XX
CC   The present invention relates to a method of producing a transformed
CC   plant that selectively increases production of gamma-aminobutyric acid
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the tobacco GAD1 protein.
XX
SQ   Sequence   496 AA;

  Query Match             85.7%;  Score 2240;  DB 23;  Length 496;
  Best Local Similarity   85.1%;  Pred. No. 1.2e-205;
  Matches  428;  Conservative   35;  Mismatches   32;  Indels    8;  Gaps    3;

Qy         1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
             ||||   ||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MVLSKTASESDVSIHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
```

```
y         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
             ||||||||||||:||:| |||||||||||||||||||||||||||||||||||| : |||||
b         61 ASFVTTWMEPECNKLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

y        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
             |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
b        121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTGANVQVCWEKFARYFEVELKE 180

y        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
             ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
b        181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

y        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
             ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||:||
b        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

y        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
             |||||||||||||||||||||||||||||||||||| |||:||||||:||  ||||||||
b        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGLEK 360

y        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
             : ||||:||: |||||||||||:| | |||||: |||:|||:|||||||||||:||||||
b        361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIIPAYTMPPNAQHVTVLRVV 420

y        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
             ||||||||||||||||||||: ||| ||:||  |:::    :|:|   :    ||:| :
b        421 IREDFSRTLAERLVIDIEKVLHELDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

y        481 QRDIITGWKKFVAD-RKKTSGIC 502
             | :| | ||||||| :|||:|:|
b        474 QLEITTAWKKFVADKKKKTNGVC 496


ESULT 5
AO15138
D   AAO15138 standard; Protein; 496 AA.
X
C   AAO15138;
X
T   25-SEP-2002  (first entry)
X
E   Tobacco GAD2.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance.
X
S   Nicotiana tabacum.
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I   Kinnersley AM,  Turano FJ;
X
R   WPI; 2002-490073/52.
R   N-PSDB; AAL43416.
X
T   Making transformed plants that selectively increase gamma-aminobutyric
T   acid production, by incorporating a DNA construct with a polynucleotide
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T   -
X
S   Claim 17; Page 58; 63pp; English.
X
C   The present invention relates to a method of producing a transformed
C   plant that selectively increases production of gamma-aminobutyric acid
```

```
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the tobacco GAD2 protein.
XX
SQ   Sequence   496 AA;

  Query Match             84.1%;  Score 2199;  DB 23;  Length 496;
  Best Local Similarity   84.1%;  Pred. No. 1.1e-201;
  Matches  423;  Conservative   34;  Mismatches   38;  Indels    8;  Gaps    3;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVLSKTASESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||: |:| |||||||||||||||||||||||||||||||||||| : |||||
Db         61 ASFVTTWMEPECNTLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPFDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
Db        181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              |||||||||||||||||||||||||||| |||||||||||||||||| |||||||||:||
Db        241 IHVDAASGGFIAPFLYPELEWDFRLPLEKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||:||||||:||  |||||:||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGIEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              : ||||:||: |||||||||||:| | |||||: |||:|||| ||||||||||:||||||
Db        361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIVLAYTMPPNAQHVTVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||||||||||||||||   :| ||:||  |:::    :|:|   :    ||:| :
Db        421 IREDFSRTLAERLVIDIEKVFHGVDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

Qy        481 QRDIITGWKKFVAD-RKKTSGIC 502
              | :| | | ||||| :|||:|:|
Db        474 QLEITTAWLKFVADKKKKTNGVC 496


RESULT 6
AAO15135
ID   AAO15135 standard; Protein; 493 AA.
XX
AC   AAO15135;
XX
DT   25-SEP-2002  (first entry)
XX
DE   A thaliana GAD4.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
```

```
   (EMER-) EMERALD BIOAGRICULTURE CORP.

   Kinnersley AM,  Turano FJ;

   WPI; 2002-490073/52.
   N-PSDB; AAL43413.

   Making transformed plants that selectively increase gamma-aminobutyric
   acid production, by incorporating a DNA construct with a polynucleotide
   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
   -

   Claim 17; Page 55-56; 63pp; English.

   The present invention relates to a method of producing a transformed
   plant that selectively increases production of gamma-aminobutyric acid
   (GABA) in response to a signal, by incorporating into the plant's genome
   a DNA construct with a non-constitutive promoter operably linked to a
   polynucleotide encoding a functional plant glutamic acid decarboxylase
   (GAD), to provide a transformed plant that expresses the GAD coding
   sequence in response to a signal. Plants of this type have an enhanced
   ability to tolerate environmental or other stresses. The present sequence
   is the A. thaliana GAD4 protein.

   Sequence   493 AA;

 Query Match             81.7%;  Score 2135.5;  DB 23;  Length 493;
 Best Local Similarity   82.4%;  Pred. No. 1.3e-195;
 Matches  411;  Conservative   32;  Mismatches   45;  Indels   11;  Gaps    2;

y          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
             ||||  |||||||:|||||||||| |||||:|||||||||||||||||||||||||||||
b          1 MVLSKTVSESDVSIHSTFASRYVRNSLPRFEMPENSIPKEAAYQIINDELMLDGNPRLNL 60

y         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
             |||||||||||||||:| ||||||||||||||||||||||||||| |||||| : | |||
b         61 ASFVTTWMEPECDKLMMESINKNYVDMDEYPVTTELQNRCVNMIARLFNAPLGDGEAAVG 120

y        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
             |||||||||||||||||||:|||||||:| | ||||||||||||||||||||||||||||
b        121 VGTVGSSEAIMLAGLAFKRQWQNKRKAQGLPYDKPNIVTGANVQVCWEKFARYFEVELKE 180

y        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
             | | | |||||| :||:|||||||||| |||||| ||||||||||||||||||:||||||
b        181 VNLREDYYVMDPVKAVEMVDENTICVAAILGSTLTGEFEDVKLLNDLLVEKNKQTGWDTP 240

y        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
             ||||||||||||||||||||||||||||||||||||||||||||||||:|| | |||:||
b        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

y        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
             |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
b        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

)y       361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
             | || ||||: ||||||||||||| | |||::  |||:|||||||||| :|||:||||||
)b       361 TGRFKIVSKENGVPLVAFSLKDSSRHNEFEVAHTLRRFGWIVPAYTMPADAQHVTVLRVV 420

Qy       421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
             |||||||||||||| | |||: ||| ||:||  |:: |:          |||:  :
Db       421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMANGKVNG----------VKKTPEET 470

Qy       481 QRDIITGWKKFVADRKKTS 499
             ||::   ||| : : |||:
Db       471 QREVTAYWKKLL-ETKKTN 488

RESULT 7
ABB91695
ID   ABB91695 standard; Protein; 493 AA.
XX
AC   ABB91695;
```

```
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 906.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP09892.
XX
PR   28-AUG-2001; 2001WO-EP09892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms -
XX
PS   Claim 5; SEQ ID NO 906; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins
CC   (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC   aligning and comparing nucleic acid or amino acid sequences from plant
CC   with nucleic acid or amino acid sequences from non-plant organisms using
CC   suitable search parameters, where plant sequences having an E-value
CC   greater by a factor of 3 than the E-value of most similar non-plant
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   493 AA;

  Query Match             81.7%;  Score 2135.5;  DB 23;  Length 493;
  Best Local Similarity   82.4%;  Pred. No. 1.3e-195;
  Matches  411;  Conservative   32;  Mismatches   45;  Indels   11;  Gaps    2;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||  |||||||:|||||||||| |||||:|||||||||||||||||||||||||||||
Db          1 MVLSKTVSESDVSIHSTFASRYVRNSLPRFEMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              |||||||||||||||:| ||||||||||||||||||||||||||| |||||| : | |||
Db         61 ASFVTTWMEPECDKLMMESINKNYVDMDEYPVTTELQNRCVNMIARLFNAPLGDGEAAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||:|||||||:| | ||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRQWQNKRKAQGLPYDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              | | | |||||| :||:|||||||||| |||||| ||||||||||||||||||:||||||
Db        181 VNLREDYYVMDPVKAVEMVDENTICVAAILGSTLTGEFEDVKLLNDLLVEKNKQTGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||:|| | |||:||
Db        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | || ||||: ||||||||||||| | |||::  |||:|||||||||| :|||:||||||
Db        361 TGRFKIVSKENGVPLVAFSLKDSSRHNEFEVAHTLRRFGWIVPAYTMPADAQHVTVLRVV 420
```

```
       421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
           |||||||||||||| | |||: ||| ||:||  |:: |:             |||:  :
       421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMANGKVNG----------VKKTPEET 470

       481 QRDIITGWKKFVADRKKTS 499
           ||::   ||| : : |||:
       471 QREVTAYWKKLL-ETKKTN 488


SULT 8
O15133
   AAO15133 standard; Protein; 494 AA.

   AAO15133;

   25-SEP-2002  (first entry)

   A thaliana GAD2.

   GAD; plant GABA production regulation; glutamic acid decarboxylase;
   plant stress; GABA; gamma-aminobutyric acid; stress resistance.

   Arabidopsis thaliana.

   WO200238736-A2.

   16-MAY-2002.
X
F  07-NOV-2001; 2001WO-US47447.
X
R  07-NOV-2000; 2000US-246367P.
X
A  (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I  Kinnersley AM,  Turano FJ;
X
R  WPI; 2002-490073/52.
R  N-PSDB; AAL43411.
X
T  Making transformed plants that selectively increase gamma-aminobutyric
T  acid production, by incorporating a DNA construct with a polynucleotide
T  encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T   -
X
S  Claim 17; Page 54; 63pp; English.
X
C  The present invention relates to a method of producing a transformed
C  plant that selectively increases production of gamma-aminobutyric acid
C  (GABA) in response to a signal, by incorporating into the plant's genome
C  a DNA construct with a non-constitutive promoter operably linked to a
C  polynucleotide encoding a functional plant glutamic acid decarboxylase
C  (GAD), to provide a transformed plant that expresses the GAD coding
C  sequence in response to a signal. Plants of this type have an enhanced
C  ability to tolerate environmental or other stresses. The present sequence
C  is the A. thaliana GAD2 protein.
XX
SQ  Sequence   494 AA;

  Query Match            81.5%;  Score 2130;  DB 23;  Length 494;
  Best Local Similarity  80.2%;  Pred. No. 4.4e-195;
  Matches  404;  Conservative   46;  Mismatches   42;  Indels   12;  Gaps    3;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              |||:   : :| || : | ||||||:||:::: ||||||:|||||| |||||||||||||
Db          1 MVLTKTAT-NDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQIIKDELMLDGNPRLNL 59

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||| ||||||||||||||||||||||||:|| ||||||||:|||||
Db         60 ASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNIIARLFNAPLEESETAVG 119

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
```

```
Db        120 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGANVQVCWEKFARYFEVELKE 179

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              | |||||||||| :| :|||||||||| |||||||||||||| ||||||:||:||||:||
Db        180 VNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVKKNEETGWNTP 239

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||:|||||||||||||||||||||||||||||||||:||  |||||||
Db        240 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRAAEDLPEEL 299

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              ||||||||||||||||||||||||:||||||||||| |||:|||||| |||:||:||:||
Db        300 IFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRLGFEGYKNVMENCIENMVVLKEGIEK 359

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              ||||||||||:|||:||||||| | | |||||:||||:|||||||||| :||||||||||
Db        360 TERFNIVSKDQGVPVVAFSLKDHSFHNEFEISEMLRRFGWIVPAYTMPADAQHITVLRVV 419

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| || ||: ||| |||::  |: :         : :   ||:  ::|
Db        420 IREDFSRTLAERLVADISKVLHELDTLPSKISKKMGI---------EGIAENVKEKKMEK 470

Qy        481 Q--RDIITGWKKFVADRKKTSGIC 502
              :   ::| ||:||| :||| :|:|
Db        471 EILMEVIVGWRKFVKERKKMNGVC 494


RESULT 9
ABB91465
ID   ABB91465 standard; Protein; 494 AA.
XX
AC   ABB91465;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 676.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP09892.
XX
PR   28-AUG-2001; 2001WO-EP09892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms -
XX
PS   Claim 5; SEQ ID NO 676; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins
CC   (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC   aligning and comparing nucleic acid or amino acid sequences from plant
CC   with nucleic acid or amino acid sequences from non-plant organisms using
CC   suitable search parameters, where plant sequences having an E-value
CC   greater by a factor of 3 than the E-value of most similar non-plant
CC   sequences are selected. The polypeptides or nucleic acids encoding them
CC   are useful for identifying modulators. The identified modulators are
CC   useful as herbicides.
XX
SQ   Sequence   494 AA;
```

```
  Query Match             81.5%;  Score 2130;  DB 23;  Length 494;
  Best Local Similarity   80.2%;  Pred. No. 4.4e-195;
  Matches  404;  Conservative   46;  Mismatches   42;  Indels   12;  Gaps    3;

            1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              |||:   : :| || : | ||||||:||:::: ||||||:|||||| |||||||||||||
            1 MVLTKTAT-NDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQIIKDELMLDGNPRLNL 59

           61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||| ||||||||||||||||||||||||:|| ||||||||:|||||
)          60 ASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNIIARLFNAPLEESETAVG 119

          121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
)         120 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGANVQVCWEKFARYFEVELKE 179

          181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              | |||||||||| :| :|||||||||| |||||||||||||| ||||||:||:||||:||
)         180 VNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVKKNEETGWNTP 239

y         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||:|||||||||||||||||||||||||||||||||:||  |||||||
)         240 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRAAEDLPEEL 299

y         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              ||||||||||||||||||||||||:||||||||||| |||:|||||| |||:||:||:||
b         300 IFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRLGFEGYKNVMENCIENMVVLKEGIEK 359

y         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              ||||||||||:|||:||||||| | | |||||:||||:|||||||||| :||||||||||
b         360 TERFNIVSKDQGVPVVAFSLKDHSFHNEFEISEMLRRFGWIVPAYTMPADAQHITVLRVV 419

y         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| || ||: ||| |||::  |: :         : :   ||:  ::|
b         420 IREDFSRTLAERLVADISKVLHELDTLPSKISKKMGI---------EGIAENVKEKKMEK 470

)y        481 Q--RDIITGWKKFVADRKKTSGIC 502
              :   ::| ||:||| :||| :|:|
)b        471 EILMEVIVGWRKFVKERKKMNGVC 494

RESULT 10
AO15134
ID   AAO15134 standard; Protein; 500 AA.
XX
AC   AAO15134;
XX
DT   25-SEP-2002  (first entry)
XX
DE   A thaliana GAD3.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   N-PSDB; AAL43412.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
PT   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT    -
XX
PS   Claim 17; Page 55; 63pp; English.
XX
CC   The present invention relates to a method of producing a transformed
CC   plant that selectively increases production of gamma-aminobutyric acid
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the A. thaliana GAD1 protein.
XX
SQ   Sequence   500 AA;

  Query Match             79.5%;  Score 2080;  DB 23;  Length 500;
  Best Local Similarity   78.8%;  Pred. No. 2.8e-190;
  Matches  395;  Conservative   44;  Mismatches   52;  Indels   10;  Gaps    1;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   |:|| |:|||||||||| |: ||::|:||||||||||||||||  ||||||||
Db          1 MVLSKTASKSDDSIHSTFASRYVRNSISRFEIPKNSIPKEAAYQIINDELKFDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              |||||||||||||||:| ||||| |:||:|||||:|||||||||| ||||||  :| |:|
Db         61 ASFVTTWMEPECDKLMMESINKNNVEMDQYPVTTDLQNRCVNMIARLFNAPLGDGEAAIG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||:|||||||||:||||||| | | |:||||||||:||| |||||||||||||
Db        121 VGTVGSSEAVMLAGLAFKRQWQNKRKALGLPYDRPNIVTGANIQVCLEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||| |||||||| :||:|||||||||  |||||| ||||||||||||||||||:||||||
Db        181 VKLREGYYVMDPDKAVEMVDENTICVVAILGSTLTGEFEDVKLLNDLLVEKNKKTGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              |||||||||||||||||:||||||||||||||||||||||||||||||:|| | |||:||
Db        241 IHVDAASGGFIAPFLYPDLEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | |||||||: ||||||||||||| | |||:::||||:|||||||||| :|||:||||||
Db        361 TGRFNIVSKENGVPLVAFSLKDSSRHNEFEVAEMLRRFGWIVPAYTMPADAQHVTVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| | |||: ||| ||:||  |:: |:             |||:  :
Db        421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMASGKVNG----------VKKTPEET 470

Qy        481 QRDIITGWKKFVADRKKTSGI 501
              ||::   |||||  :   :|:
Db        471 QREVTAYWKKFVDTKTDKNGV 491

RESULT 11
ABB91694
ID   ABB91694 standard; Protein; 500 AA.
XX
AC   ABB91694;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 905.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
```

```
  WO200210210-A2.

  07-FEB-2002.

  28-AUG-2001; 2001WO-EP09892.

  28-AUG-2001; 2001WO-EP09892.

  (FARB ) BAYER AG.

  Tietjen K,  Weidler M;

  WPI; 2002-269010/31.

  Identifying plant target proteins for herbicidally active compounds,
  comprising aligning and comparing nucleic acid or amino acid sequences
  from plant with nucleic acid or amino acid sequences from non-plant
  organisms -

  Claim 5; SEQ ID NO 905; 261pp + Sequence Listing; English.

  The invention relates to identifying target proteins
  (ABB90790-ABB94016) for herbicidally active compounds, comprising
  aligning and comparing nucleic acid or amino acid sequences from plant
  with nucleic acid or amino acid sequences from non-plant organisms using
  suitable search parameters, where plant sequences having an E-value
  greater by a factor of 3 than the E-value of most similar non-plant
  sequences are selected. The polypeptides or nucleic acids encoding them
  are useful for identifying modulators. The identified modulators are
  useful as herbicides.

  Sequence   500 AA;

Query Match             79.5%;  Score 2080;  DB 23;  Length 500;
Best Local Similarity   78.8%;  Pred. No. 2.8e-190;
Matches  395;  Conservative   44;  Mismatches   52;  Indels   10;  Gaps    1;

        1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
          ||||   |:|| |:|||||||||| |: ||::|:||||||||||||||||  ||||||||
        1 MVLSKTASKSDDSIHSTFASRYVRNSISRFEIPKNSIPKEAAYQIINDELKFDGNPRLNL 60

       61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
          |||||||||||||||:| ||||| |:||:|||||:|||||||||| |||||| : | |:|
       61 ASFVTTWMEPECDKLMMESINKNNVEMDQYPVTTDLQNRCVNMIARLFNAPLGDGEAAIG 120

      121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
          |||||||||:|||||||||:||||||| | | |:||||||||:||| |||||||||||||
      121 VGTVGSSEAVMLAGLAFKRQWQNKRKALGLPYDRPNIVTGANIQVCLEKFARYFEVELKE 180

      181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
          ||| |||||||| :||:|||||||||  |||||| ||||||||||||||||||:||||||
      181 VKLREGYYVMDPDKAVEMVDENTICVVAILGSTLTGEFEDVKLLNDLLVEKNKKTGWDTP 240

      241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
          |||||||||||||||||:||||||||||||||||||||||||||||||:|| | |||:||
      241 IHVDAASGGFIAPFLYPDLEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

      301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
          |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
      301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

      361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
          | |||||||: ||||||||||||| | |||:::||||:|||||||||| :|||:||||||
      361 TGRFNIVSKENGVPLVAFSLKDSSRHNEFEVAEMLRRFGWIVPAYTMPADAQHVTVLRVV 420

      421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
          |||||||||||||| | |||: ||| ||:||  |:: |:             |||:  :
      421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMASGKVNG----------VKKTPEET 470

      481 QRDIITGWKKFVADRKKTSGI 501
          ||::   |||||  :   :|:
      471 QREVTAYWKKFVDTKTDKNGV 491
```

```
RESULT 12
AAO15140
ID   AAO15140 standard; Protein; 502 AA.
XX
AC   AAO15140;
XX
DT   25-SEP-2002  (first entry)
XX
DE   Tomato GAD protein.
XX
KW   GAD; plant GABA production regulation; glutamic acid decarboxylase;
KW   plant stress; GABA; gamma-aminobutyric acid; stress resistance.
XX
OS   Lycopersicon esculentum.
XX
PN   WO200238736-A2.
XX
PD   16-MAY-2002.
XX
PF   07-NOV-2001; 2001WO-US47447.
XX
PR   07-NOV-2000; 2000US-246367P.
XX
PA   (EMER-) EMERALD BIOAGRICULTURE CORP.
XX
PI   Kinnersley AM,  Turano FJ;
XX
DR   WPI; 2002-490073/52.
DR   N-PSDB; AAL43418.
XX
PT   Making transformed plants that selectively increase gamma-aminobutyric
PT   acid production, by incorporating a DNA construct with a polynucleotide
PT   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
PT    -
XX
PS   Claim 17; Page 59; 63pp; English.
XX
CC   The present invention relates to a method of producing a transformed
CC   plant that selectively increases production of gamma-aminobutyric acid
CC   (GABA) in response to a signal, by incorporating into the plant's genome
CC   a DNA construct with a non-constitutive promoter operably linked to a
CC   polynucleotide encoding a functional plant glutamic acid decarboxylase
CC   (GAD), to provide a transformed plant that expresses the GAD coding
CC   sequence in response to a signal. Plants of this type have an enhanced
CC   ability to tolerate environmental or other stresses. The present sequence
CC   is the tomato GAD protein.
XX
SQ   Sequence   502 AA;

  Query Match             77.0%;  Score 2014;  DB 23;  Length 502;
  Best Local Similarity   76.0%;  Pred. No. 6e-184;
  Matches  387;  Conservative   53;  Mismatches   55;  Indels   14;  Gaps    5;

Qy          1 MVL-SHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLN 59
              ||| : :: :|: |:| |||||||:  ||:||||: |:||||||||:|||||||||||||
Db          1 MVLTTTSIRDSEESLHCTFASRYVQEPLPKFKMPKKSMPKEAAYQIVNDELMLDGNPRLN 60

Qy         60 LASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAV 119
              |||||:||||||||||||||||||||||||||||||||||||||:||||:||: : ||||
Db         61 LASFVSTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMLAHLFHAPVGDDETAV 120

Qy        120 GVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELK 179
              |||||||||||||||||||||||:|||||||| |||||||||||||||||||||||||||
Db        121 GVGTVGSSEAIMLAGLAFKRKWQSKRKAEGKPFDKPNIVTGANVQVCWEKFARYFEVELK 180

Qy        180 EVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDT 239
              |||| |||||||| :||::||||||||| |||||| ||||||||||:|| :|||||||:|
Db        181 EVKLKEGYYVMDPAKAVEIVDENTICVAAILGSTLTGEFEDVKLLNELLTKKNKETGWET 240

Qy        240 PIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEE 299
              ||||||||||||||||:|:||||||||||||||||||||||||||:||||||:|||||:|
Db        241 PIHVDAASGGFIAPFLWPDLEWDFRLPLVKSINVSGHKYGLVYAGVGWVIWRSKEDLPDE 300
```

```
          300 LIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLE 359
              |:|||||||:||||||||||||| |:||||||||||| |||:|||:||  |  || ||:
          301 LVFHINYLGSDQPTFTLNFSKGSYQIIAQYYQLIRLGFEGYKNVMKNCLSNAKVLTEGIT 360

          360 KTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRV 419
              |  ||:||||| |||:|||||:||| :| ||:|: |||:|||||||||||:|:|| ||||
          361 KMGRFDIVSKDVGVPVVAFSLRDSSKYTVFEVSEHLRRFGWIVPAYTMPPDAEHIAVLRV 420

          420 VIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKI--SLGQEKSESNSDNL----MVTV 473
              ||||||| :|||||| ||||:: |||  | |:  |      :|  :   | |    | ||
          421 VIREDFSHSLAERLVSDIEKILSELDTQPPRLPTKAVRVTAEEVRDDKGDGLHHFHMDTV 480

          474 KKSDIDKQRDIITGWKKFVADRKKTSGIC 502
                   : |:|||  |:|     |||||:|
          481 -----ETQKDIIKHWRKIAG--KKTSGVC 502

SULT 13
O15136
D   AAO15136 standard; Protein; 494 AA.
X
C   AAO15136;
X
T   25-SEP-2002  (first entry)
X
E   A thaliana GAD5.
X
W   GAD; plant GABA production regulation; glutamic acid decarboxylase;
W   plant stress; GABA; gamma-aminobutyric acid; stress resistance.
X
S   Arabidopsis thaliana.
X
N   WO200238736-A2.
X
D   16-MAY-2002.
X
F   07-NOV-2001; 2001WO-US47447.
X
R   07-NOV-2000; 2000US-246367P.
X
A   (EMER-) EMERALD BIOAGRICULTURE CORP.
X
I   Kinnersley AM,  Turano FJ;
X
R   WPI; 2002-490073/52.
R   N-PSDB; AAL43414.
X
T   Making transformed plants that selectively increase gamma-aminobutyric
T   acid production, by incorporating a DNA construct with a polynucleotide
T   encoding a plant glutamic acid decarboxylase enzyme into plant's genome
T    -
XX
PS  Claim 17; Page 56; 63pp; English.
XX
CC  The present invention relates to a method of producing a transformed
CC  plant that selectively increases production of gamma-aminobutyric acid
CC  (GABA) in response to a signal, by incorporating into the plant's genome
CC  a DNA construct with a non-constitutive promoter operably linked to a
CC  polynucleotide encoding a functional plant glutamic acid decarboxylase
CC  (GAD), to provide a transformed plant that expresses the GAD coding
CC  sequence in response to a signal. Plants of this type have an enhanced
CC  ability to tolerate environmental or other stresses. The present sequence
CC  is the A. thaliana GAD1 protein.
XX
SQ  Sequence   494 AA;

  Query Match             75.3%;  Score 1969;  DB 23;  Length 494;
  Best Local Similarity   75.2%;  Pred. No. 1.2e-179;
  Matches  373;  Conservative   56;  Mismatches   61;  Indels    6;  Gaps    3;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              |||:   |:||  :||||||||||  :||||||:: :||:||||:|||||||||||||||
Db          1 MVLA-TNSDSDEHLHSTFASRYVRAVVPRFKMPDHCMPKDAAYQVINDELMLDGNPRLNL 59

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||| |:|||||||||||||||||||||||||:||:||: | | |:|
Db         60 ASFVTTWMEPECDKLIMDSVNKNYVDMDEYPVTTELQNRCVNMIANLFHAPVGEDEAAIG 119

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
               |||||||||||||||||||||::|||:| |:||||||||||||||||||||||||||||
Db        120 CGTVGSSEAIMLAGLAFKRKWQHRRKAQGLPIDKPNIVTGANVQVCWEKFARYFEVELKE 179

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||| |||||| :||:|||||||||| |||||| ||||||| ||||| ||| ||||:||
Db        180 VKLSEDYYVMDPAKAVEMVDENTICVAAILGSTLTGEFEDVKQLNDLLAEKNAETGWETP 239

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              |||||||||||||||||:|||||||| |||||||||||||||||:|||:|| |:||||||
Db        240 IHVDAASGGFIAPFLYPDLEWDFRLPWVKSINVSGHKYGLVYAGVGWVVWRTKDDLPEEL 299

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              :|||||||||||||||||||||||:|||||| |||| |||:|:|||| :|   ||||:|
Db        300 VFHINYLGADQPTFTLNFSKGSSQIIAQYYQFIRLGFEGYKNIMENCMDNARRLREGIEM 359

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | :||||||| ||||||||||||| || |||:: ||::|||:|||||| :|||| |||||
Db        360 TGKFNIVSKDIGVPLVAFSLKDSSKHTVFEIAESLRKFGWIIPAYTMPADAQHIAVLRVV 419

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||| ||:||:  | :|::|:: ||||: |  :     :  : |:  | || : :
Db        420 IREDFSRGLADRLITHIIQVLKEIEGLPSRIAHLAA----AAAVSGDDEEVKVKTAKMSL 475

Qy        481 QRDIITGWKKFVADRK 496
              : ||   ||: |  ::
Db        476 E-DITKYWKRLVEHKR 490


RESULT 14
ABB92352
ID   ABB92352 standard; Protein; 494 AA.
XX
AC   ABB92352;
XX
DT   31-MAY-2002  (first entry)
XX
DE   Herbicidally active polypeptide SEQ ID NO 1563.
XX
KW   Herbicidal; plant; agriculture; herbicide.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200210210-A2.
XX
PD   07-FEB-2002.
XX
PF   28-AUG-2001; 2001WO-EP09892.
XX
PR   28-AUG-2001; 2001WO-EP09892.
XX
PA   (FARB ) BAYER AG.
XX
PI   Tietjen K,  Weidler M;
XX
DR   WPI; 2002-269010/31.
XX
PT   Identifying plant target proteins for herbicidally active compounds,
PT   comprising aligning and comparing nucleic acid or amino acid sequences
PT   from plant with nucleic acid or amino acid sequences from non-plant
PT   organisms -
XX
PS   Claim 5; SEQ ID NO 1563; 261pp + Sequence Listing; English.
XX
CC   The invention relates to identifying target proteins
CC   (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC   aligning and comparing nucleic acid or amino acid sequences from plant
```

```
   with nucleic acid or amino acid sequences from non-plant organisms using
   suitable search parameters, where plant sequences having an E-value
   greater by a factor of 3 than the E-value of most similar non-plant
   sequences are selected. The polypeptides or nucleic acids encoding them
   are useful for identifying modulators. The identified modulators are
   useful as herbicides.

   Sequence   494 AA;

 Query Match            75.3%;  Score 1969;  DB 23;  Length 494;
 Best Local Similarity   75.2%;  Pred. No. 1.2e-179;
 Matches  373;  Conservative   56;  Mismatches   61;  Indels    6;  Gaps    3;

              1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
                |||:   |:||  :||||||||||  :||||||:: :||:||||:|||||||||||||||
)             1 MVLA-TNSDSDEHLHSTFASRYVRAVVPRFKMPDHCMPKDAAYQVINDELMLDGNPRLNL 59

/            61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
                ||||||||||||||||| |:|||||||||||||||||||||||||:||:||: | | |:|
)            60 ASFVTTWMEPECDKLIMDSVNKNYVDMDEYPVTTELQNRCVNMIANLFHAPVGEDEAAIG 119

/           121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
                 |||||||||||||||||||||::|||:| |:||||||||||||||||||||||||||||
c           120 CGTVGSSEAIMLAGLAFKRKWQHRRKAQGLPIDKPNIVTGANVQVCWEKFARYFEVELKE 179

y           181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
                ||||| |||||| :||:|||||||||| |||||| ||||||| ||||| ||| ||||:||
b           180 VKLSEDYYVMDPAKAVEMVDENTICVAAILGSTLTGEFEDVKQLNDLLAEKNAETGWETP 239

y           241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
                |||||||||||||||||:|||||||| |||||||||||||||||:|||:|| |:||||||
b           240 IHVDAASGGFIAPFLYPDLEWDFRLPWVKSINVSGHKYGLVYAGVGWVVWRTKDDLPEEL 299

y           301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
                :|||||||||||||||||||||||:|||||| |||| |||:|:|||| :|   ||||:|
b           300 VFHINYLGADQPTFTLNFSKGSSQIIAQYYQFIRLGFEGYKNIMENCMDNARRLREGIEM 359

)y          361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
                | :||||||| ||||||||||||| || |||:: ||::|||:|||||| :|||| |||||
)b          360 TGKFNIVSKDIGVPLVAFSLKDSSKHTVFEIAESLRKFGWIIPAYTMPADAQHIAVLRVV 419

Qy          421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
                ||||||| ||:||:  | :|::|:: ||||: | :      :  : |:  | || : :
Db          420 IREDFSRGLADRLITHIIQVLKEIEGLPSRIAHLAA----AAAVSGDDEEVKVKTAKMSL 475

Qy          481 QRDIITGWKKFVADRK 496
                : ||   ||: |  ::
Db          476 E-DITKYWKRLVEHKR 490


RESULT 15
AAG13764
ID   AAG13764 standard; Protein; 494 AA.
XX
AC   AAG13764;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 13378.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
```

```
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
```

```
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             75.1%;  Score 1965;  DB 21;  Length 494;
  Best Local Similarity   75.0%;  Pred. No. 2.9e-179;
  Matches  372;  Conservative   56;  Mismatches   62;  Indels    6;  Gaps    3;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              |||:   |:||  :||||||||||  :||||||:: :||:||||:|||||||||||||||
Db          1 MVLA-TNSDSDEHLHSTFASRYVRAVVPRFKMPDHCMPKDAAYQVINDELMLDGNPRLNL 59

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||| |:|||||||||||||||||||||||||: |:||: | | |:|
Db         60 ASFVTTWMEPECDKLIMDSVNKNYVDMDEYPVTTELQNRCVNMIANFFHAPVGEDEAAIG 119

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
               |||||||||||||||||||||::|||:| |:||||||||||||||||||||||||||||
Db        120 CGTVGSSEAIMLAGLAFKRKWQHRRKAQGLPIDKPNIVTGANVQVCWEKFARYFEVELKE 179

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||| |||||| :||:|||||||||| |||||| ||||||| ||||| ||| ||||:||
Db        180 VKLSEDYYVMDPAKAVEMVDENTICVAAILGSTLTGEFEDVKQLNDLLAEKNAETGWETP 239

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              |||||||||||||||||:||||||||| |||||||||||||||||:|||:|| |:||||||
Db        240 IHVDAASGGFIAPFLYPDLEWDFRLPWVKSINVSGHKYGLVYAGVGWVVWRTKDDLPEEL 299

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              :|||||||||||||||||||||||:|||||| |||| |||:|:|||| :|   ||||:|
Db        300 VFHINYLGADQPTFTLNFSKGSSQIIAQYYQFIRLGFEGYKNIMENCMDNARRLREGIEM 359

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | :||||||| ||||||||||||| || |||:: ||::|||:|||||| :|||| |||||
Db        360 TGKFNIVSKDIGVPLVAFSLKDSSKHTVFEIAESLRKFGWIIPAYTMPADAQHIAVLRVV 419

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              ||||||| ||:||:  | :|::|:: ||||: |  :     :  : |:  | || : :
Db        420 IREDFSRGLADRLITHIIQVLKEIEGLPSRIAHLAA----AAAVSGDDEEVKVKTAKMSL 475

Qy        481 QRDIITGWKKFVADRK 496
              : ||   ||: |  ::
Db        476 E-DITKYWKRLVEHKR 490
```

```
arch completed: October 22, 2003, 15:05:26
b time : 71 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2003  Compugen Ltd.

M protein - protein search, using sw model

un on:         October 22, 2003, 14:57:03 ; Search time 44 Seconds
                                         (without alignments)
                                         1097.197 Million cell updates/sec

itle:          US-10-006-852-2
erfect score:  2615
equence:       1 MVLSHAVSESDVSVHSTFAS..........DIITGWKKFVADRKKTSGIC 502

coring table:  BLOSUM62
               Gapop 10.0 , Gapext 0.5

earched:       283308 seqs, 96168682 residues

otal number of hits satisfying chosen parameters:      283308

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

atabase :       PIR_76:*
               1:  pir1:*
               2:  pir2:*
               3:  pir3:*
               4:  pir4:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
esult            Query
 No.    Score   Match Length DB  ID                     Description
--------------------------------------------------------------------------
   1     2250    86.0    500  2  A48767                 glutamate decarbox
   2     2199    84.1    496  2  T01962                 glutamate decarbox
   3   2135.5    81.7    493  2  H84431                 probable glutamate
   4     2130    81.5    494  2  H96683                 hypothetical prote
   5     2080    79.5    500  2  G84431                 probable glutamate
   6     2014    77.0    502  2  S56177                 probable glutamate
   7     1318    50.4    514  2  JC7915                 glutamate decarbox
   8     1318    50.4    520  2  T49478                 probable glutamate
   9   1251.5    47.9    467  2  S75150                 glutamate decarbox
  10     1110    42.4    460  2  F70975                 probable glutamate
  11     1108    42.4    475  2  T36342                 probable glutamate
  12     1028    39.3    466  2  B86786                 glutamate decarbox
  13     1013    38.7    466  1  B43332                 glutamate decarbox
  14     1013    38.7    466  1  S24234                 glutamate decarbox
  15     1013    38.7    466  2  B90891                 glutamate decarbox
  16     1013    38.7    466  2  G85726                 glutamate decarbox
  17     1009    38.6    466  2  E91178                 glutamate decarbox
  18     1009    38.6    466  2  F86024                 glutamate decarbox
  19      963    36.8    467  2  AC1748                 glutamate decarbox
  20    955.5    36.5    462  2  AH1130                 glutamate decarbox
  21    953.5    36.5    467  2  AB1379                 glutamate decarbox
  22      948    36.3    464  2  AB1740                 glutamate decarbox
  23      940    35.9    464  2  AC1370                 glutamate decarbox
  24    887.5    33.9    585  2  S53072                 glutamate decarbox
  25    635.5    24.3    304  2  AE3623                 glutamate decarbox
  26    335.5    12.8    167  2  AF3623                 glutamate decarbox
  27      325    12.4    589  2  S70123                 sphingosine-1-phos
  28      314    12.0    384  2  G75102                 group II decarboxy
  29      312    11.9    363  2  E69015                 glutamate decarbox
  30      311    11.9    367  2  C69500                 group II decarboxy
  31      305    11.7    383  2  D71084                 hypothetical prote
  32      288    11.0    552  2  T33760                 sphingosine-1-phos
  33      270    10.3    544  2  C86405                 probable sphingosi
  34    269.5    10.3    488  2  B69415                 group II decarboxy
  35    264.5    10.1    542  2  T29835                 sphingosine-1-phos
  36    249.5     9.5    355  2  D84192                 glutamate decarbox
  37    246.5     9.4    568  2  JC5923                 sphingosine-1-phos
  38    243.5     9.3    473  2  G72753                 hypothetical prote
  39      225     8.6    396  2  B64306                 hypothetical prote
  40    206.5     7.9    414  2  G69536                 group II decarboxy
  41      203     7.8    454  2  G72452                 probable glutamate
  42      147     5.6    557  2  F82481                 decarboxylase, gro
  43    144.5     5.5    548  2  H82234                 probable glutamate
  44    142.5     5.4    482  2  E96500                 probable histidine
  45      137     5.2    413  1  S39554                 histidine decarbox
```

```
                                ALIGNMENTS

RESULT 1
A48767
glutamate decarboxylase (EC 4.1.1.15), calmodulin-binding - garden petunia
C;Species: Petunia x hybrida (garden petunia)
C;Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 22-Jun-1999
C;Accession: A48767
R;Baum, G.; Chen, Y.; Arazi, T.; Takatsuji, H.; Fromm, H.
J. Biol. Chem. 268, 19610-19617, 1993
A;Title: A plant glutamate decarboxylase containing a calmodulin binding domain. Cloning
A;Reference number: A48767; MUID:93374956; PMID:8366104
A;Accession: A48767
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-500 <BAU>
A;Cross-references: GB:L16797; NID:g294111; PIDN:AAA33709.1; PID:g294112
C;Superfamily: Escherichia coli glutamate decarboxylase
C;Keywords: calmodulin binding; carbon-carbon lyase; carboxy-lyase

  Query Match             86.0%;  Score 2250;  DB 2;  Length 500;
  Best Local Similarity   84.9%;  Pred. No. 1.5e-150;
  Matches  427;  Conservative   39;  Mismatches   33;  Indels    4;  Gaps    2;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||  ||:||||:|||||||||||||||||||:||||||||||||||||||||||||||
Db          1 MVLSKTVSQSDVSIHSTFASRYVRTSLPRFKMPDNSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              |||||||||||||||:| |||||||||||||||||||||||||||||||||||: |||||
Db         61 ASFVTTWMEPECDKLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEDGETAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPCDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||||||||||||::||:|||||||||| |||||||||||||| |||||||||||||||||
Db        181 VKLSEGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVEKNKETGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||:|||||||||||||||||||||||||||||||||:||||:|||:||
Db        241 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRNKDDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              ||||||||||||||||||||||||||||||||||||:|||:||||||:||  ||||||||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGYEGYKNVMENCQENASVLREGLEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | ||||:||: |||||||||||:  | |||||: |||:||||||||||||||||||||||
Db        361 TGRFNIISKEIGVPLVAFSLKDNRQHNEFEISETLRRFGWIVPAYTMPPNAQHITVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| |||||: ||| ||:||  |::: :|:: :|   :    ||:| :
```

```
b        421 IREDFSRTLAERLVRDIEKVLHELDTLPARVNAKLAVAEEQAAANGSEVH---KKTDSEV 477

y        481 QRDIITGWKKFVAD-RKKTSGIC 502
             | ::|| ||||| : :|||: :|
b        478 QLEMITAWKKFVEEKKKKTNRVC 500


ESULT 2
01962
lutamate decarboxylase (EC 4.1.1.15) 2, calmodulin-binding - common tobacco
,Species: Nicotiana tabacum (common tobacco)
,Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 22-Jun-1999
,Accession: T01962
,Yun, S.J.; Oh, S.H.
ol. Cells 8, 125-129, 1998
,Title: Cloning and characterization of a tobacco cDNA encoding calcium/calmodulin-depe
,Reference number: Z14473; MUID:98302498; PMID:9638642
,Accession: T01962
,Status: preliminary; translated from GB/EMBL/DDBJ
,Molecule type: mRNA
 Residues: 1-496 <YUN>
 Cross-references: EMBL:AF020424; NID:g3252853; PIDN:AAC39483.1; PID:g3252854
 Genetics:
 Gene: GAD2
 Function:
 Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
 Superfamily: Escherichia coli glutamate decarboxylase
 Keywords: calmodulin binding; carbon-carbon lyase; carboxy-lyase

 Query Match             84.1%;  Score 2199;  DB 2;  Length 496;
 Best Local Similarity   84.1%;  Pred. No. 5.8e-147;
 Matches  423;  Conservative   34;  Mismatches   38;  Indels    8;  Gaps    3;

           1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
             ||||   |||||||||||||||||||||||||||||||||||||||||||||||||||||
           1 MVLSKTASESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60

          61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
             ||||||||||||: |:| |||||||||||||||||||||||||||||||||| : |||||
          61 ASFVTTWMEPECNTLMMDSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLGDGETAVG 120

         121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
             |||||||||||||||||||||||| ||:||| ||||||||||||||||||||||||||||
         121 VGTVGSSEAIMLAGLAFKRKWQNKMKAQGKPFDKPNIVTGANVQVCWEKFARYFEVELKE 180

         181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
             ||||:|||||||::||:|||||||||| |||||||||||||| |||||:|||||||||||
         181 VKLSDGYYVMDPEKAVEMVDENTICVAAILGSTLNGEFEDVKRLNDLLIEKNKETGWDTP 240

         241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
             ||||||||||||||||||||||||||| ||||||||||||||||||| |||||||||:||
         241 IHVDAASGGFIAPFLYPELEWDFRLPLEKSINVSGHKYGLVYAGIGWAIWRNKEDLPDEL 300

         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
             |||||||||||||||||||||||||||||||||||| |||:||||||:||  |||||:||
         301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYKNVMENCQENARVLREGIEK 360

         361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
             : ||||:||: |||||||||||:| | |||||: |||:|||| ||||||||||:||||||
         361 SGRFNIISKEIGVPLVAFSLKDNSQHNEFEISETLRRFGWIVLAYTMPPNAQHVTVLRVV 420

         421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
             ||||||||||||||||||||   :| ||:||  |:::    :|:|   :     ||:| :
         421 IREDFSRTLAERLVIDIEKVFHGVDTLPARVNAKLAV----AEANGSGVH---KKTDREV 473

         481 QRDIITGWKKFVAD-RKKTSGIC 502
             | :| | | |||||  :|||:|:|
         474 QLEITTAWLKFVADKKKKTNGVC 496


SULT 3
4431
obable glutamate decarboxylase [imported] - Arabidopsis thaliana
```

```
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: H84431
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.,
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84431
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-493 <STO>
A;Cross-references: GB:AE002093; NID:g4406789; PIDN:AAD20099.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g02010
A;Map position: 2
C;Superfamily: Escherichia coli glutamate decarboxylase

  Query Match             81.7%;  Score 2135.5;  DB 2;  Length 493;
  Best Local Similarity   82.4%;  Pred. No. 1.7e-142;
  Matches  411;  Conservative   32;  Mismatches   45;  Indels   11;  Gaps    2;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||  |||||||:|||||||||| |||||:|||||||||||||||||||||||||||||
Db          1 MVLSKTVSESDVSIHSTFASRYVRNSLPRFEMPENSIPKEAAYQIINDELMLDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              |||||||||||||||:| ||||||||||||||||||||||||||| |||||| : | |||
Db         61 ASFVTTWMEPECDKLMMESINKNYVDMDEYPVTTELQNRCVNMIARLFNAPLGDGEAAVG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||||||||||||:|||||||:| | ||||||||||||||||||||||||||||
Db        121 VGTVGSSEAIMLAGLAFKRQWQNKRKAQGLPYDKPNIVTGANVQVCWEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              | | | |||||| :||:|||||||||| |||||| |||||||||||||||||:|||||||
Db        181 VNLREDYYVMDPVKAVEMVDENTICVAAILGSTLTGEFEDVKLLNDLLVEKNKQTGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||:|| | |||:||
Db        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | || ||||: ||||||||||||| | |||::  |||:|||||||||| :|||:||||||
Db        361 TGRFKIVSKENGVPLVAFSLKDSSRHNEFEVAHTLRRFGWIVPAYTMPADAQHVTVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| | |||: ||| ||:||  |:: |:          |||:  :
Db        421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMANGKVNG----------VKKTPEET 470

Qy        481 QRDIITGWKKFVADRKKTS 499
              ||::   ||| : : |||:
Db        471 QREVTAYWKKLL-ETKKTN 488


RESULT 4
H96683
hypothetical protein F12P19.12 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C;Accession: H96683
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
```

```
,Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
er, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
,Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
,Reference number: A86141; MUID:21016719; PMID:11130712
,Accession: H96683
,Status: preliminary
,Molecule type: DNA
,Residues: 1-494 <STO>
,Cross-references: GB:AE005173; NID:g6227020; PIDN:AAF06056.1; GSPDB:GN00141
,Genetics:
,Gene: F12P19.12
,Map position: 1
,Superfamily: Escherichia coli glutamate decarboxylase

 Query Match             81.5%;  Score 2130;  DB 2;  Length 494;
 Best Local Similarity   80.2%;  Pred. No. 4.1e-142;
 Matches  404;  Conservative   46;  Mismatches   42;  Indels   12;  Gaps    3;

            1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              |||:   : :| || : | ||||||:||:::: ||||||:|||||| |||||||||||||
            1 MVLTKTAT-NDESVCTMFGSRYVRTTLPKYEIGENSIPKDAAYQIIKDELMLDGNPRLNL 59

           61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              ||||||||||||||||| ||||||||||||||||||||||||:|| ||||||||:|||||
           60 ASFVTTWMEPECDKLIMDSINKNYVDMDEYPVTTELQNRCVNIIARLFNAPLEESETAVG 119

          121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
          120 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPYDKPNIVTGANVQVCWEKFARYFEVELKE 179

          181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              | |||||||||| :| :|||||||||| |||||||||||||| ||||||:||:||||:||
          180 VNLSEGYYVMDPDKAAEMVDENTICVAAILGSTLNGEFEDVKRLNDLLVKKNEETGWNTP 239

          241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              ||||||||||||||:|||||||||||||||||||||||||||||||||:||  |||||||
          240 IHVDAASGGFIAPFIYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRAAEDLPEEL 299

          301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              ||||||||||||||||||||||||:||||||||||| |||:|||||| |||:||:||:||
          300 IFHINYLGADQPTFTLNFSKGSSQIIAQYYQLIRLGFEGYKNVMENCIENMVVLKEGIEK 359

          361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              ||||||||||:|||:||||||| | | |||||:||||:|||||||||| :||||||||||
          360 TERFNIVSKDQGVPVVAFSLKDHSFHNEFEISEMLRRFGWIVPAYTMPADAQHITVLRVV 419

          421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| || ||: ||| |||::  |: :         : :   ||:  ::|
          420 IREDFSRTLAERLVADISKVLHELDTLPSKISKKMGI---------EGIAENVKEKKMEK 470

          481 Q--RDIITGWKKFVADRKKTSGIC 502
              :   ::| ||:||| :||| :|:|
          471 EILMEVIVGWRKFVKERKKMNGVC 494


SULT 5
4431
obable glutamate decarboxylase [imported] - Arabidopsis thaliana
Species: Arabidopsis thaliana (mouse-ear cress)
Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
Accession: G84431
Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
ss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
ture 402, 761-768, 1999
Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
Reference number: A84420; MUID:20083487; PMID:10617197
Accession: G84431
Status: preliminary
Molecule type: DNA
Residues: 1-500 <STO>
Cross-references: GB:AE002093; NID:g4406783; PIDN:AAD20093.1; GSPDB:GN00139
Genetics:
```

```
A;Gene: At2g02000
A;Map position: 2
C;Superfamily: Escherichia coli glutamate decarboxylase

  Query Match             79.5%;  Score 2080;  DB 2;  Length 500;
  Best Local Similarity   78.8%;  Pred. No. 1.4e-138;
  Matches  395;  Conservative   44;  Mismatches   52;  Indels   10;  Gaps    1;

Qy          1 MVLSHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNL 60
              ||||   |:|| |:|||||||||| |: ||::|:||||||||||||||||  ||||||||
Db          1 MVLSKTASKSDDSIHSTFASRYVRNSISRFEIPKNSIPKEAAYQIINDELKFDGNPRLNL 60

Qy         61 ASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVG 120
              |||||||||||||||:| ||||| |:||:|||||:|||||||||| |||||| : | |:|
Db         61 ASFVTTWMEPECDKLMMESINKNNVEMDQYPVTTDLQNRCVNMIARLFNAPLGDGEAAIG 120

Qy        121 VGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
              |||||||||:|||||||||:||||||| | | |:||||||||:||| |||||||||||||
Db        121 VGTVGSSEAVMLAGLAFKRQWQNKRKALGLPYDRPNIVTGANIQVCLEKFARYFEVELKE 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              ||| |||||||| :||:|||||||||  |||||| ||||||||||||||||||:||||||
Db        181 VKLREGYYVMDPDKAVEMVDENTICVVAILGSTLTGEFEDVKLLNDLLVEKNKKTGWDTP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              |||||||||||||||||:||||||||||||||||||||||||||||||:|| | |||:||
Db        241 IHVDAASGGFIAPFLYPDLEWDFRLPLVKSINVSGHKYGLVYAGIGWVVWRTKTDLPDEL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              |||||||||||||||||||||||||||||||||||| |||||||:||||||:|||:||||
Db        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGFEGYRNVMDNCRENMMVLRQGLEK 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
              | |||||||: ||||||||||||| | |||:::||||:|||||||||| :|||:||||||
Db        361 TGRFNIVSKENGVPLVAFSLKDSSRHNEFEVAEMLRRFGWIVPAYTMPADAQHVTVLRVV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDELPSRVIHKISLGQEKSESNSDNLMVTVKKSDIDK 480
              |||||||||||||| | |||: ||| ||:||  |:: |:             |||:  : 
Db        421 IREDFSRTLAERLVADFEKVLHELDTLPARVHAKMASGKVNG----------VKKTPEET 470

Qy        481 QRDIITGWKKFVADRKKTSGI 501
              ||::   |||||  :   :|:
Db        471 QREVTAYWKKFVDTKTDKNGV 491

RESULT 6
S56177
probable glutamate decarboxylase - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 28-May-1999
C;Accession: S56177
R;Gallego, P.P.; Whotton, L.; Picton, S.; Grierson, D.; Gray, J.E.
Plant Mol. Biol. 27, 1143-1151, 1995
A;Title: A role for glutamate decarboxylase during tomato ripening: the characterisatior
A;Reference number: S56177; MUID:95284363; PMID:7766895
A;Accession: S56177
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-502 <GAL>
A;Cross-references: GB:X80840; NID:g993002; PIDN:CAA56812.1; PID:g995555
C;Superfamily: Escherichia coli glutamate decarboxylase

  Query Match             77.0%;  Score 2014;  DB 2;  Length 502;
  Best Local Similarity   76.0%;  Pred. No. 6.1e-134;
  Matches  387;  Conservative   53;  Mismatches   55;  Indels   14;  Gaps    5;

Qy          1 MVL-SHAVSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLN 59
              ||| : :: :|: |:| |:| |||||||:  ||:||||: |:||||||||:|||||||||||||
Db          1 MVLTTTSIRDSEESLHCTFASRYVQEPLPKFKMPKKSMPKEAAYQIVNDELMLDGNPRLN 60

Qy         60 LASFVTTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAV 119
              |||||:||||||||||||||||||||||||||||||||||:||||:||: : ||||
Db         61 LASFVSTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMLAHLFHAPVGDDETAV 120
```

```
y           120 GVGTVGSSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELK 179
                |||||||||||||||||||||||:|||||||| |||||||||||||||||||||||||||
b           121 GVGTVGSSEAIMLAGLAFKRKWQSKRKAEGKPFDKPNIVTGANVQVCWEKFARYFEVELK 180

y           180 EVKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDT 239
                |||| |||||||| :||::||||||||| |||||| ||||||||||:|| :|||||||:|
b           181 EVKLKEGYYVMDPAKAVEIVDENTICVAAILGSTLTGEFEDVKLLNELLTKKNKETGWET 240

y           240 PIHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEE 299
                ||||||||||||||||:|:||||||||||||||||||||||||||:||||||:|||||:|
b           241 PIHVDAASGGFIAPFLWPDLEWDFRLPLVKSINVSGHKYGLVYAGVGWVIWRSKEDLPDE 300

y           300 LIFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLE 359
                |:|||||||:||||||||||||| |:||||||||||| |||:|||:||  |  || ||:
b           301 LVFHINYLGSDQPTFTLNFSKGSYQIIAQYYQLIRLGFEGYKNVMKNCLSNAKVLTEGIT 360

y           360 KTERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRV 419
                |  ||:||||| |||:|||||:||| :| ||:|: |||:|||||||||||:|:|| ||||
b           361 KMGRFDIVSKDVGVPVVAFSLRDSSKYTVFEVSEHLRRFGWIVPAYTMPPDAEHIAVLRV 420

y           420 VIREDFSRTLAERLVIDIEKVMRELDELPSRVIHKI--SLGQEKSESNSDNL----MVTV 473
                ||||||| :|||||| ||||:: |||  | |:  |      :|  :   | |    | ||
b           421 VIREDFSHSLAERLVSDIEKILSELDTQPPRLPTKAVRVTAEEVRDDKGDGLHHFHMDTV 480

y           474 KKSDIDKQRDIITGWKKFVADRKKTSGIC 502
                     :  |:|||  |:|     |||||:|
b           481 -----ETQKDIIKHWRKIAG--KKTSGVC 502


ESULT 7
7915
lutamate decarboxylase (EC 4.1.1.15) - Aspergillus oryzae
 Species: Aspergillus oryzae
 Date: 31-Mar-2003 #sequence_revision 31-Mar-2003 #text_change 21-Apr-2003
 Accession: JC7915
 Kato, Y., Kato, Y., Furukawa, K., and Hara, S.
iosci. Biotechnol. Biochem. 66, 2600-2605, 2002
 Title: Cloning and nucleotide sequence of the glutamate decarboxylase-encoding gene ga
;Reference number: JC7915; MUID:22484577; PMID:12596854
 Accession: JC7915
 Molecule type: DNA
 Residues: 1-514 <KAT>
 Cross-references: DDBJ:AB025422
 Comment: This enzyme produces gamma-aminobutyric acid.
 Genetics:
 Gene: gadA
 Introns: 79/2; 88/3; 477/2
 Keywords: carbon-carbon lyase; carboxy-lyase

 Query Match             50.4%;  Score 1318;  DB 2;  Length 514;
 Best Local Similarity   56.6%;  Pred. No. 6e-85;
 Matches  256;  Conservative   77;  Mismatches  109;  Indels   10;  Gaps    5;

y             8 SESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTW 67
                :: |    | : :|:    ||: :|||  :|:| ||::| ||| ||||| |||||||||:
b            28 TDEDGFYSSVYGTRFATEQLPQTEMPEREMPREVAYRMIKDELSLDGNPMLNLASFVTTY 87

y            68 MEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPL-EEAETAVGVGTVGS 126
                || | :||:  | :||::| :||| : |:||||||||| |||||:  | |  :|  |:||
b            88 MEDEAEKLMTESFSKNFIDYEEYPQSAEIQNRCVNMIARLFNAPVHSEDEHPMGTSTIGS 147

y           127 SEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEG 186
                ||||||  || ||:||||||||||   :||||  : ||||||| ||||:|| : |  :|
b           148 SEAIMLGTLAMKRRWQNKRKAEGKDYSRPNIVMNSAVQVCWEKAARYFDVEERYVYCTED 207

y           187 YYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAA 246
                 ||:|||||||:||||||  :  |||:|  ||:|||| :||||:|:|     | |||||||
b           208 RYVIDPQQAVDLVDENTIGICAILGTTYTGEYEDVKAINDLLIERN----IDVPIHVDAA 263

y           247 SGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINY 306
                ||||:|||: |:|||||||| | ||||||||||||| |:|||:||: | ||::|||:|||
b           264 SGGFVAPFINPKLEWDFRLPKVVSINVSGHKYGLVYPGVGWVVWRSPEYLPKDLIFNINY 323
```

```
Qy          307 LGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNI 366
                |||:| :||||||||:|||| ||||:||||  |||::| |       | : |||   | |
Db          324 LGAEQASFTLNFSKGASQVIGQYYQMIRLGKRGYRSIMTNITVTADFLAQELEKM-GFII 382

Qy          367 VS--KDEGVPLVAFSL--KDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIR 422
                :|  :  |:||||| |  :      || ::  ||  ||||||||| ||: :: ::|||:|
Db          383 MSQRRGHGLPLVAFRLPAEREGQFDEFALAHQLRERGWIVPAYTMAPNSNNLKLMRVVVR 442

Qy          423 EDFSRTLAERLVIDIEKVMRELDELPSRVIHK 454
                |||:::  : |: ||:  :: | ::   :: |
Db          443 EDFTKSRCDALLSDIKLGLKTLGDMDKAMLDK 474


RESULT 8
T49478
probable glutamate decarboxylase [imported] - Neurospora crassa
N;Alternate names: protein B14D6.300
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000
C;Accession: T49478
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49478
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-520 <SCH>
A;Cross-references: EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.300
A;Experimental source: BAC clone B14D6; strain OR74A
C;Genetics:
A;Gene: NCSP:B14D6.300
A;Map position: 6
A;Introns: 85/2; 504/3
C;Superfamily: Escherichia coli glutamate decarboxylase

  Query Match             50.4%;  Score 1318;  DB 2;  Length 520;
  Best Local Similarity   55.6%;  Pred. No. 6.1e-85;
  Matches  262;  Conservative   73;  Mismatches  118;  Indels   18;  Gaps    6;

Qy            8 SESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTW 67
                :: |    | : |::    ||| :||:  :||| ||::| ||| ||||| |||||||||:
Db           34 NDEDRFTTSVYGSKFAAADLPRHEMPDEEMPKEVAYRMIKDELSLDGNPLLNLASFVTTY 93

Qy           68 MEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAE--TAVGVGTVG 125
                || | :||:  |: ||::| :||| | ::|||||:||  |||||:::||  :|||  :||
Db           94 MEEEAEKLMTESLPKNFIDYEEYPQTADIQNRCVSMIGRLFNAPVKDAEASSAVGTSSVG 153

Qy          126 SSEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSE 185
                |||||||  || |::|:||| ||||| ||||::  : |||||||  |||||| | |  :
Db          154 SSEAIMLGVLAMKKRWKNKRIAEGKPADKPNLIMSSAVQVCWEKATRYFEVEEKFVYCTP 213

Qy          186 GYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDA 245
                  ||:||::  ||:|||||| :  |||:|  ||:|||| :||||||:    | ||||||||
Db          214 DRYVIDPKETVDLVDENTIGICCILGTTYTGEYEDVKAVNDLLVER----GLDTPIHVDA 269

Qy          246 ASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHIN 305
                |||||:|||: |:|||||||  | ||||||||||||| |:|||:||: | ||:||:|:||
Db          270 ASGGFVAPFVVPDLEWDFRLKNVVSINVSGHKYGLVYPGVGWVVWRSAEYLPQELVFNIN 329

Qy          306 YLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFN 365
                |||||| :||||||||:|||| |||||||||  ||| :| |       | | |     |
Db          330 YLGADQASFTLNFSKGASQVIGQYYQLIRLGKHGYRAIMSNLTRTADYLAESLAAL-GFI 388

Qy          366 IVSK--DEGVPLVAFSLKDSSCHT--EFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVI 421
                |:|:   :|:||||| ||:    |  || ::  ||  ||||||||| |  : : :||:|:
Db          389 IMSQKSGQGLPLVAFRLKEDPDRTYDEFALAHQLRVRGWIVPAYTMAPKTEGLKMLRIVV 448

Qy          422 REDFSRTLAERLVIDI-------EKVMRELDELPSRVIHKISLGQEKSESN 465
                ||||||   : |: ||       |:: :|  :     |||  :   |:  |
Db          449 REDFSRNRCDGLISDIRSCQGILEQMDKETVKKQQEFIHKHHVVSGKASHN 499
```

```
ESULT 9
75150
lutamate decarboxylase - Synechocystis sp. (strain PCC 6803)
,Alternate names: protein sll1641
,Species: Synechocystis sp.
,Variety: PCC 6803
,Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
,Accession: S75150
,Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
NA Res. 3, 109-136, 1996
,Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

,Reference number: S74322; MUID:97061201; PMID:8905231
,Accession: S75150
,Status: nucleic acid sequence not shown; translation not shown
,Molecule type: DNA
,Residues: 1-467 <KAN>
,Cross-references: EMBL:D90903; GB:AB001339; NID:g1652127; PIDN:BAA17064.1; PID:g165214
,Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
,Genetics:
,Gene: gad
,Superfamily: Escherichia coli glutamate decarboxylase

 Query Match            47.9%;  Score 1251.5;  DB 2;  Length 467;
 Best Local Similarity  50.7%;  Pred. No. 2.5e-80;
 Matches  229;  Conservative   89;  Mismatches  125;  Indels    9;  Gaps    3;

y          7 VSESDVSVHSTFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTT 66
             :||::  :  |:|:| :  |: :::|||  :    || :|:||| |||| |||||:||||
b         11 LSEAESLLTPTYAARGLANSVSKYEMPETEMLPAIAYNLIHDELGLDGNSRLNLATFVTT 70

y         67 WMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGS 126
             |||||  :|:  : :|| :| |||| | |:: ||||::: |:|||   :  | |  |:||
b         71 WMEPEARQLMADTFDKNMIDKDEYPQTAEIELRCVNILSRLWNAP--ASAEATGCSTIGS 128

y        127 SEAIMLAGLAFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEG 186
             ||| || |:| | ||: :|:| ||| |:||:| | ||||||||| ||:||| : | :
b        129 SEAAMLGGMAMKWKWRQRRQAAGKPGDRPNLVMGINVQVCWEKFCRYWEVEPRFVPMEGD 188

y        187 YYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAA 246
              | : |::|| ::||||| |  ||||| :| :| :: ||| |   |: |||  |:|:|||
b        189 RYHISPEEAVKLIDENTIGVIGILGSTFDGSYEPIEALNDALETLNQRTGWQVPLHIDAA 248

y        247 SGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINY 306
             ||||||||| |:| |||||| ||||| ||||||||| |:||:|||:||:|||||||| ||
b        249 SGGFIAPFLDPDLRWDFRLPWVKSINTSGHKYGLVYPGVGWIIWRDKEELPEELIFHCNY 308

y        307 LGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNI 366
             || | | ||||:  :||:||||  :||| |||| : : ||:  : |   : :   | :
b        309 LGGDLPNFALNFSRPGNQVVAQYYNFLRLGKEGYRKIQQTCRDTALYLSGKIAQLGPFEL 368

y        367 VSKDEGVPLVAFSLKD----SSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIR 422
             ::    :|: |: |||    ::|:| ::::| ||  ||:|||| || | : : | |:|::
b        369 LTDGGDIPVFAWRLKDEVLANTCYTLYDMADKLRERGWLVPAYRMPKNREDLVVQRIVVK 428

y        423 EDFSRTLAERLVIDIEKVMRELDELPSRVIHK 454
             | ||| :|: |: |:|: :      |    ||
b        429 EGFSRDMADLLLADMERAIAYFASQPD---HK 457


ESULT 10
70975
robable glutamate decarboxylase - Mycobacterium tuberculosis (strain H37RV)
,Species: Mycobacterium tuberculosis
,Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
,Accession: F70975
,Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
ajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
ature 393, 537-544, 1998
 Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
```

```
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: F70975
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-460 <COL>
A;Cross-references: GB:Z95389; GB:AL123456; NID:g3242261; PIDN:CAB08681.1; PID:g2104353
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: gadB
C;Superfamily: Escherichia coli glutamate decarboxylase

  Query Match            42.4%;  Score 1110;  DB 2;  Length 460;
  Best Local Similarity  48.1%;  Pred. No. 2.2e-70;
  Matches  215;  Conservative   77;  Mismatches  145;  Indels   10;  Gaps    4;

Qy          8 SESDVSVHS---TFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFV 64
              |   |  ||    :  |     :|  :||: |:  ||||: |:|||||||: |||||:||
Db          4 SHPSVPAHSIAPAYTGRMFTAPVPALRMPDESMDPEAAYRFIHDELMLDGSSRLNLATFV 63

Qy         65 TTWMEPECDKLIMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAP---LEEAETAVGV 121
              ||||:|| :||:  : :|| :| |||| |  :: |||:|:| ||:|      :  :| ||
Db         64 TTWMDPEAEKLMAETFDKNMIDKDEYPATAAIEARCVSMVADLFHAEGLRDHDPTSATGV 123

Qy        122 GTVGSSEAIMLAGLAFKRKWQNKRKA-EGKPVDKPNIVTGANVQVCWEKFARYFEVELKE 180
               |:|||||:|| ||| | :|: :  : :|:    ||:| |:|||| |||| |||:|| :
Db        124 STIGSSEAVMLGGLALKWRWRQRVGSWKGR---MPNLVMGSNVQVVWEKFCRYFDVEPRY 180

Qy        181 VKLSEGYYVMDPQQAVDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTP 240
              : :  | ||: |:| :  |||||| |  |||:|  || | :  :   | :     | | |
Db        181 LPMERGRYVITPEQVLAAVDENTIGVVAILGTTYTGELEPIAEICAALDKLAAGGGVDVP 240

Qy        241 IHVDAASGGFIAPFLYPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEEL 300
              :|||||||||: |||:|:| |||||| | ||||||||||| | |:|:|:||  | |||:|
Db        241 VHVDAASGGFVVPFLHPDLVWDFRLPRVVSINVSGHKYGLTYPGVGFVVWRGPEHLPEDL 300

Qy        301 IFHINYLGADQPTFTLNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEK 360
              :| :|||| | ||||||||:  :||: |||  :||| :||  ||:        | : | :
Db        301 VFRVNYLGGDMPTFTLNFSRPGNQVVGQYYNFLRLGRDGYTKVMQALSHTARWLGDQLRE 360

Qy        361 TERFNIVSKDEGVPLVAFSLKDSSCHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVV 420
               :   ::|    :|:|:| |     :|||::|  || :|| ||||||| ||  : |||:|
Db        361 VDHCEVISDGSAIPVVSFRLAGDRGYTEFDVSHELRTFGWQVPAYTMPDNATDVAVLRIV 420

Qy        421 IREDFSRTLAERLVIDIEKVMRELDEL 447
              :||  |  ||  |  |    :  ||::
Db        421 VREGLSADLARALHDDAVTALAALDKV 447


RESULT 11
T36342
probable glutamate decarboxylase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C;Accession: T36342
R;Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z21575
A;Accession: T36342
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-475 <OLI>
A;Cross-references: EMBL:AL049841; PIDN:CAB42769.1; GSPDB:GN00070; SCOEDB:SCE9.23
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: gad; SCOEDB:SCE9.23
C;Superfamily: Escherichia coli glutamate decarboxylase

  Query Match            42.4%;  Score 1108;  DB 2;  Length 475;
  Best Local Similarity  49.8%;  Pred. No. 3.2e-70;
  Matches  211;  Conservative   79;  Mismatches  128;  Indels    6;  Gaps    4;

Qy         28 PRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIMSSINKNYVDM 87
              |  ::|:: :| |:||::::|||||||| |||||:||||||||:   |:    :|| :|
```

```
b           40 PAHRLPDSPLPPESAYRLVHDELMLDGNARLNLATFVTTWMEPQAGVLMSECRDKNMIDK 99

y           88 DEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAFKRKW--QNKR 145
               |||| | ||: ||| |:| |::||  :  ||||  | ||||| ||||:| ||:|  :|
b          100 DEYPRTAELERRCVAMLADLWHAP--DPSTAVGCSTTGSSEACMLAGMALKRRWALRNAD 157

y          146 KAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQAVDMVDENTIC 205
               :   | | :||:| | ||||||:||  ::||| ::| :    : :||| | :: |||||
b          158 RYPAKDV-RPNLVMGVNVQVCWDKFCNFWEVEARQVPMEGDRFHLDPQAAAELCDENTIG 216

y          206 VADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFLYPELEWDFRL 265
               |  ||||| :| :| |  |   |    : || | |:||| |||  :||||  :| |||||
b          217 VVGILGSTFDGSYEPVADLCAALDALQERTGLDVPVHVDGASGAMVAPFLDEDLVWDFRL 276

y          266 PLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFTLNFSKGSSQV 325
               | | ||| ||||||||| |:|| :||: | |||||:| :|||| | ||| ||||:  :||
b          277 PRVASINTSGHKYGLVYPGVGWALWRDAEALPEELVFRVNYLGGDMPTFALNFSRPGAQV 336

y          326 IAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPLVAFSLKDS-S 384
               :||||  :||| |||| | :: |:    | | :     | :::: : :|: ||:  |  :
b          337 VAQYYNFLRLGREGYRAVQQSARDIAGSLAERVAALGDFRLLTRGDQLPVFAFTTADDVT 396

y          385 CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERLVIDIEKVMREL 444
                :  |::|  ||  ||:||||| ||: : ::||||| |  ||  :|: |: |:|::: ||
b          397 AYDVFDVSRRLREGGWLVPAYTFPPHREDLSVLRVVCRNGFSADMADLLLADLERLLPEL 456

y          445 DELP 448
                  |
b          457 RRQP 460


ESULT 12
86786
lutamate decarboxylase (EC 4.1.1.15) [imported] - Lactococcus lactis subsp. lactis (str
,Species: Lactococcus lactis subsp. lactis
,Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
,Accession: B86786
,Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
enome Res. 11, 731-753, 2001
,Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
,Reference number: A86625; MUID:21235186; PMID:11337471
,Accession: B86786
,Status: preliminary
 Molecule type: DNA
 Residues: 1-466 <STO>
 Cross-references: GB:AE005176; PID:g12724267; PIDN:AAK05388.1; GSPDB:GN00146
 Experimental source: strain IL1403
 Genetics:
 Gene: gadB
 Superfamily: Escherichia coli glutamate decarboxylase
 Keywords: carbon-carbon lyase; carboxy-lyase

 Query Match             39.3%;  Score 1028;  DB 2;  Length 466;
 Best Local Similarity   45.0%;  Pred. No. 1.3e-64;
 Matches  195;  Conservative   91;  Mismatches  141;  Indels    6;  Gaps    3;

y           18 FASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKLIM 77
               | |   :  ||::|: : ||    |||:: ||:: :|| |||||:|  |:||||  ||:
b           18 FGSESEQVDLPKYKLAQQSIEPRVAYQLVQDEMLDEGNARLNLATFCQTYMEPEAVKLMS 77

y           78 SSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGLAF 137
                ::  || :|  ||| |||::|||||||| |:||   | |  :|  |:||||| || |:|
b           78 QTLEKNAIDKSEYPRTTEIENRCVNMIADLWNA--SEKEKFMGTSTIGSSEACMLGGMAM 135

y          138 KRKWQNKRKAEGKPVD--KPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
               |  |: : :  |  ::  |||:|  :  |||||||  |:::|::|| : : :  ::  :
b          136 KFSWRKRAEKLGLDINAKKPNLVISSGYQVCWEKFCIYWDIEMREVPMDKEHMSINLDKV 195

y          196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
               :| ||| || |  |:| |  | ::|:| |::|: | ||:| :   |||||||||  |||:
b          196 MDYVDEYTIGVVGIMGITYTGRYDDIKALDNLIEEYNKQTDYKVYIHVDAASGGLYAPFV 255

y          256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
                |||||||||  | ||| ||||||||| |:|||:||:|: ||||||| ::||| : ||
Db         256 EPELEWDFRLKNVISINTSGHKYGLVYPGVGWVLWRDKKYLPEELIFKVSYLGGELPTMA 315

Qy         316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIVSKDEGVPL 375
               :|||  :||:| |||  :| | :||: : |   :  : | : :|||  | |::    :|:
Db         316 INFSHSASQLIGQYYNFVRYGFDGYKAIHERTHKVAMFLAKEIEKTGMFEIMNDGSQLPI 375

Qy         376 VAFSLKDSSCH--TEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLAERL 433
               | : ||: |      ::::| |   || |||| :| | ::  : |:||| ||   :|
Db         376 VCYKLKEDSNRGWNLYDLADRLLMKGWQVPAYPLPKNLENEIIQRLVIRADFGMNMAFNY 435

Qy         434 VIDIEKVMRELDE 446
               | |::: :  |::
Db         436 VQDMQEAIEALNK 448


RESULT 13
B43332
glutamate decarboxylase (EC 4.1.1.15) beta - Escherichia coli (strain K-12)
N;Alternate names: L-glutamate 1-carboxy-lyase
C;Species: Escherichia coli
C;Date: 10-Mar-1994 #sequence_revision 23-Mar-1995 #text_change 01-Mar-2002
C;Accession: B43332; S30261; H64902
R;Smith, D.K.; Kassam, T.; Singh, B.; Elliott, J.F.
J. Bacteriol. 174, 5820-5826, 1992
A;Title: Escherichia coli has two homologous glutamate decarboxylase genes that map to
A;Reference number: A43332; MUID:92394884; PMID:1522060
A;Accession: B43332
A;Molecule type: DNA
A;Residues: 1-466 <SMI>
A;Cross-references: GB:M84025; NID:g146059; PIDN:AAA23834.1; PID:g146060
R;Yoshida, T.; Ueguchi, C.; Yamada, H.; Mizuno, T.
Mol. Gen. Genet. 237, 113-122, 1993
A;Title: Function of the Escherichia coli nucleoid protein, H-NS: molecular analysis of
A;Reference number: S30261; MUID:93204884; PMID:8455549
A;Accession: S30261
A;Molecule type: protein
A;Residues: 1-4,'LQVL',7-15 <YOS>
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64902
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-466 <BLAT>
A;Cross-references: GB:AE000246; GB:U00096; NID:g1787764; PIDN:AAC74566.1; PID:g1787769,
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: gadB
A;Map position: 78 min
C;Function:
A;Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
A;Note: in E. coli, two isoforms (alpha and beta) have been found, each encoded by a se
C;Superfamily: Escherichia coli glutamate decarboxylase
C;Keywords: carbon-carbon lyase; carboxy-lyase; homohexamer; phosphoprotein; pyridoxal
F;276/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

  Query Match             38.7%;  Score 1013;  DB 1;  Length 466;
  Best Local Similarity   45.0%;  Pred. No. 1.5e-63;
  Matches  197;  Conservative   83;  Mismatches  152;  Indels    6;  Gaps    3;

Qy          16 STFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKL 75
               | | :: : |     : | : :  : |:||||||| |||| | |||:|  || :    ||
Db          16 SRFGAKSISTIAESKRFPLHEMRDDVAFQIINDELYLDGNARQNLATFCQTWDDENVHKL 75

Qy          76 IMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGL 135
               :  |||||::| :||| :  :  |||||:| |::||  :   |||  |:||||| || |:
Db          76 MDLSINKNWIDKEEYPQSAAIDLRCVNMVADLWHAPAPKNGQAVGTNTIGSSEACMLGGM 135

Qy         136 AFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
               | | :|: : :| ||| ||||:| |  ||:|| |||||::|||:|: :  |   |||::
```

```
b         136 AMKWRWRKRMEAAGKPTDKPNLVCGP-VQICWHKFARYWDVELREIPMRPGQLFMDPKRM 194

y         196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
              ::  ||||| |    | |  | :|  : |:| | :   :|| |  :|:|||||||:|||:
b         195 IEACDENTIGVVPTFGVTYTGNYEFPQPLHDALDKFQADTGIDIDMHIDAASGGFLAPFV 254

y         256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
               |:: |||||| ||||: ||||:||   | ||||||::| ||:||:|:::|||    ||
b         255 APDIVWDFRLPRVKSISASGHKFGLAPLGCGWVIWRDEEALPQELVFNVDYLGGQIGTFA 314

y         316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIV---SKDEG 372
              :|||: : |||||||: :||| |||  |     :    | : : |   :  :     |||
b         315 INFSRPAGQVIAQYYEFLRLGREGYTKVQNASYQVAAYLADEIAKLGPYEFICTGRPDEG 374

y         373 VPLVAFSLKDSS--CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLA 430
              :| | | |||     :| :::|: ||  || |||:|:   |  | |:|:: |  |    |
b         375 IPAVCFKLKDGEDPGYTLYDLSERLRLRGWQVPAFTLGGEATDIVVMRIMCRRGFEMDFA 434

y         431 ERLVIDIEKVMRELDELP 448
              | |: | :  :: | : |
o         435 ELLLEDYKASLKYLSDHP 452


ESULT 14
24234
lutamate decarboxylase (EC 4.1.1.15) alpha - Escherichia coli (strain K-12)
,Alternate names: GAD alpha protein; L-glutamate 1-carboxy-lyase
,Species: Escherichia coli
,Date: 22-Jan-1993 #sequence_revision 05-Dec-1997 #text_change 01-Mar-2002
,Accession: S47737; H65149; S24234; S23421; A43332; PN0616
,Plunkett, G.
ubmitted to the EMBL Data Library, March 1994
,Reference number: S47666
,Accession: S47737
,Status: preliminary
,Molecule type: DNA
,Residues: 1-466 <PLU>
,Cross-references: EMBL:U00039; NID:g466582; PIDN:AAB18493.1; PID:g466654
,Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
cience 277, 1453-1462, 1997
 Title: The complete genome sequence of Escherichia coli K-12.
,Reference number: A64720; MUID:97426617; PMID:9278503
 Accession: H65149
 Status: nucleic acid sequence not shown; translation not shown
 Molecule type: DNA
 Residues: 1-466 <BLAT>
 Cross-references: GB:AE000428; GB:U00096; NID:g1789931; PIDN:AAC76542.1; PID:g1789934;
 Experimental source: strain K-12, substrain MG1655
 Maras, B.; Sweeney, G.; Barra, D.; Bossa, F.; John, R.A.
ır. J. Biochem. 204, 93-98, 1992
 Title: The amino acid sequence of glutamate decarboxylase from Escherichia coli. Evolu
 Reference number: S23421; MUID:92155241; PMID:1740158
 Accession: S24234
 Molecule type: protein
 Residues: 1-63,'S',65-72,'R',74-152,'N',154-164,'S',166-354,'N',356-466 <MAR>
 Accession: S23421
 Molecule type: DNA
 Residues: 148-164,'S',166-207,'N',209-294,'V',296-431 <MAR1>
 Cross-references: GB:X63123; NID:g41601; PIDN:CAA44834.1; PID:g938166
 Smith, D.K.; Kassam, T.; Singh, B.; Elliott, J.F.
  Bacteriol. 174, 5820-5826, 1992
,Title: Escherichia coli has two homologous glutamate decarboxylase genes that map to d
,Reference number: A43332; MUID:92394884; PMID:1522060
,Accession: A43332
,Molecule type: DNA
,Residues: 1-466 <SMI>
 Cross-references: GB:M84024; NID:g146057; PIDN:AAA23833.1; PID:g146058
 Note: sequence extracted from NCBI backbone (NCBIN:112979, NCBIP:112980)
 Yoshida, T.; Yamashino, T.; Ueguchi, C.; Mizuno, T.
ıosci. Biotechnol. Biochem. 57, 1568-1569, 1993
 Title: Expression of the Escherichia coli dimorphic glutamic acid decarboxylases is re
 Reference number: PN0616; MUID:94033862; PMID:7764225
```

```
A;Accession: PN0616
A;Molecule type: protein
A;Residues: 1-21,'X' <YOS>
C;Genetics:
A;Gene: gadA
A;Map position: 78 min
C;Function:
A;Description: catalyzes the formation of gamma-aminobutyric acid, from L-glutamic acid
A;Note: in E. coli, two isoforms (alpha and beta) have been found, each encoded by a sep
C;Superfamily: Escherichia coli glutamate decarboxylase
C;Keywords: carbon-carbon lyase; carboxy-lyase; homohexamer; phosphoprotein; pyridoxal
F;276/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

  Query Match             38.7%;  Score 1013;  DB 1;  Length 466;
  Best Local Similarity   45.0%;  Pred. No. 1.5e-63;
  Matches  197;  Conservative   83;  Mismatches  152;  Indels    6;  Gaps    3;

Qy         16 STFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKL 75
              | | :: : |     : | : :  : |:||||||| |||| | |||:|  || :    ||
Db         16 SRFGAKAISTIAESKRFPLHEMRDDVAFQIINDELYLDGNARQNLATFCQTWDDENVHKL 75

Qy         76 IMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGL 135
              :  |||||::| :||| :  :  |||||:| |::||  :   |||  |:||||| || |:
Db         76 MDLSINKNWIDKEEYPQSAAIDLRCVNMVADLWHAPAPKNGQAVGTNTIGSSEACMLGGM 135

Qy        136 AFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
              | |:|: : :| ||| ||||:| |  ||:|| |||||::|||:|: :  |   |||::
Db        136 AMKWRWRKRMEAAGKPTDKPNLVCGP-VQICWHKFARYWDVELREIPMRPGQLFMDPKRM 194

Qy        196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
              ::  ||||| |    | |  | :|  : |:| | :   :|| |  :|:|||||||:|||:
Db        195 IEACDENTIGVVPTFGVTYTGNYEFPQPLHDALDKFQADTGIDIDMHIDAASGGFLAPFV 254

Qy        256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
               |:: |||||| ||||: ||||:||   | ||||||::| ||:||:|:::|||    ||
Db        255 APDIVWDFRLPRVKSISASGHKFGLAPLGCGWVIWRDEEALPQELVFNVDYLGGQIGTFA 314

Qy        316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIV---SKDEG 372
              :|||: : |||||||: :||| |||  |     :    | : : |   :  :     |||
Db        315 INFSRPAGQVIAQYYEFLRLGREGYTKVQNASYQVAAYLADEIAKLGPYEFICTGRPDEG 374

Qy        373 VPLVAFSLKDSS--CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLA 430
              :| | | |||     :| :::|: ||  || |||:|:   |  | |:|:: |  |    |
Db        375 IPAVCFKLKDGEDPGYTLYDLSERLRLRGWQVPAFTLGGEATDIVVMRIMCRRGFEMDFA 434

Qy        431 ERLVIDIEKVMRELDELP 448
              | |: | :  :: | : |
Db        435 ELLLEDYKASLKYLSDHP 452


RESULT 15
B90891
glutamate decarboxylase isozyme [imported] - Escherichia coli (strain O157:H7, substrai
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: B90891
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90891
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB35521.1; PID:g13361564; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2098
C;Superfamily: Escherichia coli glutamate decarboxylase

  Query Match             38.7%;  Score 1013;  DB 2;  Length 466;
  Best Local Similarity   45.0%;  Pred. No. 1.5e-63;
```

```
 Matches  197;  Conservative   83;  Mismatches  152;  Indels    6;  Gaps    3;

y        16 STFASRYVRTSLPRFKMPENSIPKEAAYQIINDELMLDGNPRLNLASFVTTWMEPECDKL 75
            | | :: : |     : | : :  : |:||||||| |||| | |||:|  || :    ||
b        16 SRFGAKSISTIAESKRFPLHEMRDDVAFQIINDELYLDGNARQNLATFCQTWDDENVHKL 75

y        76 IMSSINKNYVDMDEYPVTTELQNRCVNMIAHLFNAPLEEAETAVGVGTVGSSEAIMLAGL 135
            :  |||||::| :||| :  :  |||||:| |::||   :  |||  |:||||| || |:
b        76 MDLSINKNWIDKEEYPQSAAIDLRCVNMVADLWHAPAPKNGQAVGTNTIGSSEACMLGGM 135

y       136 AFKRKWQNKRKAEGKPVDKPNIVTGANVQVCWEKFARYFEVELKEVKLSEGYYVMDPQQA 195
            | | :|: :  | ||| ||||:| |  ||:|| |||||::|||:|: :  |   |||::
b       136 AMKWRWRKRMEAAGKPTDKPNLVCGP-VQICWHKFARYWDVELREIPMRPGQLFMDPKRM 194

y       196 VDMVDENTICVADILGSTLNGEFEDVKLLNDLLVEKNKETGWDTPIHVDAASGGFIAPFL 255
            ::  ||||| |    | |  | :|  : |:| | :   :|| |  :|:|||||||:|||:
b       195 IEACDENTIGVVPTFGVTYTGNYEFPQPLHDALDKFQADTGIDIDMHIDAASGGFLAPFV 254

y       256 YPELEWDFRLPLVKSINVSGHKYGLVYAGIGWVIWRNKEDLPEELIFHINYLGADQPTFT 315
             |:: |||||| ||||: ||||:||   | ||||||::| ||:||:|:::|||    ||
b       255 APDIVWDFRLPRVKSISASGHKFGLAPLGCGWVIWRDEEALPQELVFNVDYLGGQIGTFA 314

y       316 LNFSKGSSQVIAQYYQLIRLGHEGYRNVMENCRENMIVLREGLEKTERFNIV---SKDEG 372
            :|||: : |||||||: :||| |||  |      :   | : : |   :  :     |||
b       315 INFSRPAGQVIAQYYEFLRLGREGYTKVQNASYQVAAYLADEIAKLGPYEFICTGRPDEG 374

y       373 VPLVAFSLKDSS--CHTEFEISDMLRRYGWIVPAYTMPPNAQHITVLRVVIREDFSRTLA 430
            :| | | |||     :| :::|: ||  || |||:|:   |  | |:|:: |  |    |
b       375 IPAVCFKLKDGEDPGYTLYDLSERLRLRGWQVPAFTLGGEATDIVVMRIMCRRGFEMDFA 434

y       431 ERLVIDIEKVMRELDELP 448
            | |: | :  :: | : |
b       435 ELLLEDYKASLKYLSDHP 452

earch completed: October 22, 2003, 15:08:36
ob time : 45 secs
```